Quality indicators for the care of osteoarthritis in general practice: identification, synthesis, and implementation

John James Edwards

Doctor of Philosophy

June 2017

Keele University

Keele University

Declaration Part 1. To be bound in the thesis

1.1 SUBMISSION OF THESIS FOR A RESEARCH DEGREE

1.1.1 Part I. DECLARATION by the candidate for a research degree. To be bound in the thesis

Degree for which thesis being submitted Doctor of Philosophy

Title of thesis Quality indicators for the care of osteoarthritis in general practice: identification, synthesis, and implementation

This thesis contains confidential information and is subject to the protocol set down for the submission and examination of such a thesis.

~~YES/~~NO [please delete as appropriate; if YES the box in Part II should be completed]

Date of submission 01 March 2017 Original registration date 27.09.2010

1. (Date of submission must comply with Regulation 2D)

Name of candidate John James EDWARDS
Research Institute Primary Care & Health Sciences
Name of Lead Supervisor Professor Kelvin Jordan

I certify that:

(a) The thesis being submitted for examination is my own account of my own research
(b) My research has been conducted ethically. Where relevant a letter from the approving body confirming that ethical approval has been given has been bound in the thesis as an Annex
(c) The data and results presented are the genuine data and results actually obtained by me during the conduct of the research
(d) Where I have drawn on the work, ideas and results of others this has been appropriately acknowledged in the thesis
(e) Where any collaboration has taken place with one or more other researchers, I have included within an 'Acknowledgments' section in the thesis a clear statement of their contributions, in line with the relevant statement in the Code of Practice (see Note overleaf).
(f) The greater portion of the work described in the thesis has been undertaken subsequent to my registration for the higher degree for which I am submitting for examination
(g) Where part of the work described in the thesis has previously been incorporated in another thesis submitted by me for a higher degree (if any), this has been identified and acknowledged in the thesis
(h) The thesis submitted is within the required word limit as specified in the Regulations

Total words in submitted thesis (including text and footnotes, but excluding references and appendices) 83,451

Signature of candidate  Date 01.03.2017

Note

Extract from Code of Practice: If the research degree is set within a broader programme of work involving a group of investigators – particularly if this programme of work predates the candidate's registration – the candidate should provide an explicit statement (in an 'Acknowledgments' section) of the respective roles of the candidate and these other individuals in relevant aspects of the work reported in the thesis. For example, it should make clear, where relevant, the candidate's role in designing the study, developing data collection instruments, collecting primary data, analysing such data, and formulating conclusions from the analysis. Others involved in these aspects of the research should be named, and their contributions relative to that of the candidate should be specified (*this does not apply to the ordinary supervision, only if the supervisor or supervisory team has had greater than usual involvement*).

Acknowledgements

I was involved with the development of the MOSAICS study protocol[1] from its inception, as part of a multidisciplinary team. I was co-principal investigator on the medical records review (consultation data and template) component of the MOSAICS study.

Elements of the work reported in this thesis have been used, in adapted form, in a peer-reviewed publication [see Appendix A, Appendix B] and in another publication currently in submission [Jordan KP, Edwards JJ, Porcheret M, et al. 2017]. All the analyses reported in this thesis were developed and conducted by me. Compared to those used in the associated publications, they are new or extend the published analyses through inclusion of additional independent variables (selected and justified by me in Chapter Four) and, for Chapters Four to Six, through additional multilevel methods such as random slope analysis, three-level analyses (patients within clinicians within practices) and estimation of the clinician-level variance partition coefficient. In Chapter Seven, the chart of rates of routinely-recorded OA management by month was produced by Professor Jordan for the associated publication [Appendix B]. Where relevant, each Chapter methods section reports the extension or differences to analyses conducted by Professor Jordan for the associated publications. The conclusions from the results are my own and substantially extend the discussion sections in the associated publications, which I led.

I wish to express my profound thanks to my supervisory team, Professor Kelvin Jordan, Professor George Peat, and Dr John Bedson, for all their support, advice and encouragement in developing, undertaking, and writing up this endeavour. Grateful thanks also go to Professor Krysia Dziedzic for all her advice and encouragement, both as a supervisor in the initial stages of this work when I was registered to undertake an MPhil, and since as an unfailing source of advice and optimism.

My thanks also go to all those who made this research possible: the staff members of the general practices that participated in the MOSAICS study and their patients; the MOSAICS study team – including Professor Krysia Dziedzic as chief investigator, the study coordinators (June Handy, Angela Pushpa-Rajah, and Sue Hill), the wider study team and trial steering committee; my co-reviewers

for the systematic review; the general practitioners who helped with the definition of clinical osteoarthritis determined by Read codes (Christian Mallen, Mark Porcheret, John Bedson, Richard Hayward, and Eyitope Roberts); the Health Informatics staff (especially Tracy Whitehurst, Simon Wathall, and Ian Thomas) for their expertise and assistance in creating the electronic recording template and in defining the data collection requirements, and obtaining the data from the practices to the data download specification.

To my parents, sister and brother, my thanks as ever for your love and confidence, without which this enterprise would never have been started! Last, but not least, my heartfelt thanks for their unfailing love and support, and seemingly endless patience, go to my wife, Barbara, and children, James and Freja.

Abstract

Background

Previous studies have demonstrated suboptimal management of care for osteoarthritis (OA). The objectives of this study were to (i) identify indicators of quality of care for OA in general practice, (ii) measure quality of care using routine general practice records and through an enhanced recording template (iii) estimate the effect of the template introduction on quality of care, and (iv) assess the feasibility of quality indicators as trial outcome measures.

Methods

A systematic review and narrative synthesis of quality indicators was undertaken. An iterative process of development resulted in an electronic template to record management of OA in consultations, based on identified quality indicators. This was triggered by a case definition of clinical OA derived through consensus. An assessment of coding, diagnostic misclassification using consultation narrative, and baseline recorded quality of care before template installation in eight practices was undertaken. Measurement after template installation facilitated a before-and-after comparison of care. The indicators were used as secondary outcomes in a cluster-randomised trial of a model OA consultation.

Results

There were fifteen valid, feasible quality indicators. Consultation prevalence of clinical OA was comparable to other estimates but up to one-third of cases may not represent true OA. Prescribing and referral data were well-captured in the routine record; assessment and core treatment indicators (such as education and advice) were not and so were included in the recording template. The template had small-to-moderate effects on weight recording, and paracetamol and topical anti-inflammatory prescription.

Assessment of the effect of the model consultation was limited by high baseline quality achievement and variation between trial arms, practices and clinicians.

Conclusion

Assessment of quality of care for OA in general practice through quality indicators is feasible but comprehensive assessment requires enhanced recording approaches. Inter-clinician variability requires further understanding and reduction, and triangulation with patient-experienced quality is needed.

Abbreviations

ACR	American College of Rheumatology
ASR	Age-Standardised Rate
BMI	Body Mass Index
BNF	British National Formulary
c/o	complaining of
CI	Confidence Interval
CiPCA	Consultations in Primary Care Archives
DALY	Disability-Adjusted Life Year
EHR	Electronic Health Record
EMIS	Egton Medical Information Systems
ESP	European Standard Population
EULAR	European League Against Rheumatism
FDA	Food & Drug Administration
GP	General Practitioner
GPRD	General Practice Research Database
HIV	Human immunodeficiency virus
ICC	Intraclass correlation coefficient
ICD-10	International Classification of Disease, version 10
IQR	Inter-quartile range
MOAC	Model OA Consultation
MOSAICS	Management of OSteoArthritis In ConsultationS
MQL1	First-order maximum quasi-likelihood
MQL2	Second-order maximum quasi-likelihood
MRI	Magnetic Resonance Imaging
NICE	National Institute for Health & Care Excellence
NSAID	Non-Steroidal Anti-Inflammatory Drug
OA	Osteoarthritis
OARSI	OsteoArthritis Research Society International
OR	Odds Ratio
PDSA	Plan-Do-Study-Act
PPV	Positive Predictive Value
PQL1	First-order penalised quasi-likelihood
PQL2	Second-order penalised quasi-likelihood
QI	Quality indicator
QOF	Quality & Outcomes Framework
QS	Quality Standard (NICE)
RCGP	Royal College of General Practitioners
RIPCHS	Research Institute for Primary Care & Health Sciences
RRR	Relative Risk Ratio
SRR	Standardised Rate Ratio
UK	United Kingdom
US	Ultrasound
USA	United States of America
VPC	Variance Partition Coefficient
WHO	World Health Organization
WRS	Weekly Returns Service (of the RCGP)
YLD	Years Lived with Disability
YLL	Years of Life Lost

Table of Contents

List of Boxes

List of Tables

List of Figures

List of Appendices

Publications and presentations arising from the work described in this thesis

Peer-reviewed publications

- Edwards JJ, Khanna M, Jordan KP, et al. Quality indicators for the primary care of osteoarthritis: a systematic review. Ann Rheum Dis 2015;74(3):490-8. [Appendix A]
- Edwards JJ, Jordan KP, Peat G, et al. Quality of care for OA: the effect of a point-of-care consultation recording template. Rheumatology (Oxford) 2015;54(5):844-53. [Appendix B]

Invited scientific session speaker presentation (EULAR Congress 2015)

- Edwards JJ. SP0237 An Overview of Quality Indicators for the Management of Osteoarthritis in Primary Care. Ann Rheum Dis 2015;74(Suppl 2):57-58.

Oral presentations

Society for Academic Primary Care (North-West) 2011

- Edwards JJ, Khanna M, Jordan KP, et al. A Systematic Review of Quality Indicators in the Primary Care of Osteoarthritis

EULAR Congress 2015

- Edwards JJ, Jordan KP, Porcheret M, et al. OP0106 Effect of a Model Consultation on Quality of Care of Osteoarthritis: A Primary Care Cluster Randomised Trial. Ann Rheum Dis 2015;74(Suppl 2):108-09.

Poster presentations

EULAR Congress 2011

- Edwards JJ, Khanna M, Jordan KP, et al. Quality Indicators for the Primary Care of Osteoarthritis: A Systematic Review. (Abstract). Ann Rheum Dis 2011;70 (Suppl3):338.

British Society for Rheumatology Conference 2014

- Edwards JJ, Jordan KP, Peat G, et al. 69. Quality of Care for Osteoarthritis: The Effect of a Point-of-Care Consultation Recording Template. Rheumatology (Oxford) 2014;53 (suppl 1):i81.

Evidence Live 2015

- Edwards JJ, Jordan KP, Jinks C, Bedson J, Clarkson K, Hay EM, Dziedzic KS. A model consultation for osteoarthritis: the effect on the recorded quality of primary care.

Chapter One: Introduction and background literature

1.1 Introduction

This thesis will describe the identification, synthesis, and implementation of quality indicators for the primary care of peripheral joint (hand, hip, knee, foot) osteoarthritis (OA). To investigate how these might be implemented in practice and in clinical trials to improve adherence to the National Institute for Health & Care Excellence* (NICE) OA management guidelines, a case definition for clinical OA in primary care records was developed and tested. Using this definition, the feasibility of assessment of quality of care for OA through the routine general practice medical record was investigated; the overcoming of important identified recording deficiencies through an enhanced recording methodology is described, with the resulting level of recorded quality of care outlined as well as an estimate of the effect of the change in record quality after implementation of the enhanced recording approach. Finally, the implementation of the quality indicators as outcome measures in the clinical trial is reported and overall conclusions presented.

For context, this chapter identifies an appropriate definition of OA and its recognition for clinical and research purposes. The impacts that OA has on populations and its clinical management as applied to general practice in England are described. The previously-reported quality of care for OA is outlined, with a description of the methods of assessment of quality, and a summary of methods for improving care for OA in general practice. The thesis aims and objectives are set out and the outline of the thesis including the programme of studies within which the research for this thesis was conducted is briefly described.

*NICE is constituted to provide guidance to England only, though some products are also supplied to the other UK home nations https://www.nice.org.uk/about/who-we-are [Accessed 21/09/2016]

1.2 Definitions of osteoarthritis

OA does not have an established, single, generally accepted definition. It has been variously defined in terms of pathological changes that occur in joints,[2] radiographic changes,[3] and symptoms.[4]

The broad concept of osteoarthritis that was proposed by NICE is used as the foundation of much of the work in this thesis. This states that:

> *"Osteoarthritis is defined not as a disease or a single condition but as a common complex disorder with multiple risk factors."* (NICE 2008, p. 3[5])

For practical clinical purposes, the concept of OA has to be converted into a clinically recognisable and relevant working definition. This could be based on findings from radiographs, clinical history and examination, or on histopathology, only the first two of which are applicable to general practice. The means of recognising OA in a patient is not always straightforward: one systematic review identified 25 different classification criteria (clinical, radiological, or combined) for knee OA alone.[6]

1.2.1 Radiographic OA

Plain film radiography has historically been the principal imaging modality used to identify OA. Classification systems and score such as the Kellgren-Lawrence scale[3] and minimal joint space[7] have been used extensively to define OA in research studies. However, the use of plain radiography to define OA, especially of the knee, has been increasingly criticised. The finding of radiographic features of OA is sensitive to positioning[8] and the number of views taken.[9] Plain X-rays also do not image the soft tissues associated with joints, which are known to contribute to OA symptomatology.[10] The European League Against Rheumatism (EULAR) recommendations for diagnosis of knee OA suggest multiple views of both knees as the gold standard for morphological assessment of knee OA through plain radiography and also suggest that further imaging modalities are not usually required for a diagnosis of OA.[11]

The contemporary role of X-rays in general practice management of OA is not clearly understood. Although NICE argue[4] that radiographic investigation is not needed for diagnosis, and that the need for onward referral is poorly correlated with radiographic findings, Bedson et al. have previously found that general practitioners' (GPs') decisions about management of knee pain are associated with a decision to request an X-ray; the decision to X-ray was thought to be a clinician characteristic rather than the result of the clinical presentation.[12] Radiographic OA has been suggested to be more likely to be reported if it had been raised as a clinical diagnosis by the clinician requesting the X-ray.[13] There is also evidence that the link between joint pain and radiological findings of OA are not consistent: for example, one study of hip pain in Framingham study participants identified a sensitivity of only 16.5% for radiological OA in patients with hip pain localised to the groin.[14] Knee pain symptoms and radiographic OA have also been found to be only weakly associated.[15]

The EULAR recommendations for the diagnosis of knee OA indicate that a clinical diagnosis of OA can be made even in the absence of radiographic changes if symptoms (activity-related pain, no more than short-lived early morning stiffness, and functional limitation) and at least one examination finding (crepitus, painful or restricted movement, bony enlargement, and no more than a modest joint effusion) are present.[11] Guidance in England also argues that X-rays do not have an important role in the diagnosis and management of OA in primary care, and so a clinical definition of OA that is not dependent upon radiological investigations should be used.[4]

1.2.2 Clinically apparent OA

As the pathological processes involved in OA are not highly correlated with clinical syndrome of pain and disability,[16] from a clinician's perspective an alternative approach to the definition and recognition of the clinical syndrome of OA is required. The 2008 National Collaborating Centre for Chronic Conditions OA guidelines recommended the use of the following clinical

working diagnosis of OA, assuming that non-OA inflammatory arthritis and connective tissue disease had been excluded:[4]

> *"The* [Guideline Development Group] *considered the following to represent a clinician's working diagnosis of peripheral joint osteoarthritis:*
> ● *persistent joint pain that is worse with use* ● *age 45 years old and over*
> ● *morning stiffness lasting no more than half an hour."* (p. 9)

Such an approach has the advantage that the diagnosis is not dependent upon further investigations. However, the definition is more open to interpretation than some other primary care diagnoses, which may be recorded on the basis of a laboratory or near-patient test, or specialist assessment. Based on this, there is the potential for substantial variability between (and possibly within) clinicians in the making and recording of a diagnosis of OA. In an analysis of influences on the decision to record an OA diagnosis, Jordan et al. concluded that some cases of recorded joint pain (which in the UK is a much more commonly recorded diagnostic term than OA) represented early OA, with some cases fitting the clinical and radiographic criteria for diagnosis.[17] This again suggests that GPs are not consistent in coding OA as such but that some cases may rather be recorded as joint pain. It is unclear if GPs manage patients on the basis of a working diagnosis of OA even where it is not formally recorded as such.

There has been an attempt to reconcile the pathological/radiological and clinical approaches, by a joint OARSI – Food and Drug Administration (FDA) initiative, which used OA 'disease' to mean the structural changes that occur within a osteoarthritic joint, and OA 'illness' to refer to the clinical symptoms experienced by patients with OA.[18] Under this definition, the work presented in this thesis is more concerned with the latter.

1.3 The population burden of OA

1.3.1 Epidemiology of OA

Estimates of the prevalence of OA in population surveys vary widely and depend on the case definition used.[19] A series of population surveys undertaken between 2002 and 2005 in North Staffordshire estimated the prevalence of OA in the hand, hip, knee or foot (the joints applicable to the Management of OSteoArthritis In ConsultationS (MOSAICS) studies, described in section 1.7, below). The overall estimate for OA affecting at least one joint was 532 per 10,000 people aged 50 and over, with 219 per 10,000 having disabling OA (equivalent to 714 per 10,000 in the general population).[20]

For research based in general practice consulters, such as the MOSAICS studies,[1] estimates of the proportion of the registered population consulting for OA during a defined period of time ('consultation prevalence') are particularly relevant. The main sources of such estimates in the UK are general practice medical record databases. Jordan et al.[21] compared rates of recorded OA between four primary care datasets. The recorded annual consultation prevalence of OA was estimated as 164 to 426 per 10,000 people aged 15 and over; joint pain was estimated at 273 to 433 per 10,000. In a second study, which used data from a local network of general practices in North Staffordshire (the Consultations in Primary Care Archives [CiPCA][21,22]), the one-year consultation prevalence for diagnosed OA was estimated at 447 per 10,000 people aged 45 years and over.[23] Using a wider definition of OA (including certain joint pain codes, further discussed in Chapter Three, section 3.5), the annual consultation prevalence estimates in people aged 45 and over increased to 1192 per 10,000; when a seven-year period was used, the consultation prevalence estimate was 3483 per 10,000.

It is difficult to estimate with accuracy the expected proportion of people registered with a GP who would be expected to have clinical OA (the 'illness' as described by the OARSI-FDA initiative) but it might be as high as 35% of adults aged 45 and over as suggested by the 7-

year period consultation prevalence reported by Jordan et al.[23]; only around one-third of these would be expected to consult in a one-year period. This estimate may be inflated, especially in relatively young patients, by people consulting with symptoms that cause little (or temporary) disability. By comparison, estimates from the North Staffordshire Osteoarthritis Project (NorStOP) suggested a 22% prevalence of knee pain associated with some disability in people aged 50 and over.[20]

1.3.2 Assessment of OA burden of disease

There are limitations in the assessment of the population burden of OA that flow from the differences in case definitions of OA, referred to above. The Global Burden of Disease study[24] based estimates on a systematic review of hip and knee OA prevalence (which itself used estimates of prevalence from radiographic OA with or without symptoms).[25] It estimated the number of years lived with disability (YLD) due to OA at 197 per 100,000 population (95% confidence interval [CI] 134–279) in 1999 and 249 per 100,000 (95% CI 172–352) in 2010, a 26.2% increase.[26] In the UK, OA was ranked as the 11th most common cause of YLDs; there were an estimated 351 per 100,000 (95% CI 221–520) disability-adjusted life-years (DALY).[27]

Costs associated with OA are substantial. The 2008 NICE guidelines reported hip and knee joint replacement costs at £405 million in the year 2000 and wider societal costs at £3.2 billion in lost production (due to working days lost) plus £43 million in community services and £215 million on social services.[4] In the 2012 *OA Nation* report, Arthritis Care estimated that at a patient level, additional OA-related costs were in the order of £480 per patient, or £2.6 billion annually across the country.[28] Overall, societal costs (in higher-income countries) have been estimated as 0.25-0.50% of gross domestic product.[29]

OA should therefore be regarded as a common condition with a substantial associated degree of ill-health and high healthcare and wider societal costs. As such, it is an important condition for primary health care services to treat as effectively as possible to reduce the burden of disease.

1.4 Clinical management of OA

Some risks for OA may be ameliorated through promotion of maintenance of a healthy weight and avoidance of trauma[30] but beyond these, there are no clearly-established preventative actions to be implemented at an individual level. The majority of clinical interventions attempt to treat symptoms once OA is established.

1.4.1 Summary of recommended care for OA in general practice

There are various international guidelines, including those from EULAR,[11,31-34] OARSI,[35-37] and the American College of Rheumatology (ACR),[38] but identification and treatment of OA in general practice in England has been guided by the recommendations made by NICE 2008,[5] updated in 2014.[39] Guidance for the management of OA has diverged in some respects, notably regarding the use of SYmptomatic Slow-Acting Drugs in OsteoArthritis (SYSADOA), as well as in the detail of oral NSAID use and gastroprotection. In general, however, there is not much heterogeneity between different sources of guidance on the management of OA.

NICE (2008) recommended a holistic assessment, core interventions of education, exercise advice and weight loss advice where relevant, followed by relatively safe pharmacological pain management options (paracetamol, topical non-steroidal anti-inflammatory drugs [NSAIDs]) and then adjuvant pharmacological (analgesics) options, non-pharmacological management (e.g. ambulatory or non-ambulatory assistive devices, orthotics, thermotherapy, manual therapy), and joint arthroplasty.[5] NICE also recommended against certain processes of care, notably topical rubefacients, glucosamine, and arthroscopic debridement and washout.

The principles of quality measurement are described below, with a summary of the general practice delivery of OA care and the extent to which clinicians actually adhere to management guidelines.

1.4.2 Measurement of quality of care

Structures, processes, outcomes, and the concept of case mix

The quality of care delivered by health services may be assessed in a number of ways. Donabedian divided quality assessment into the domains of structures (such as the premises from which healthcare is provided, equipment, or staffing levels), processes (the care delivered to patients, including diagnosis and treatments) and outcomes (the end points most relevant to patients such as health status or quality of life).[40,41] Patient experience has been added to these assessment domains, with evidence that patient experience is linked to the quality of care delivered (clinical effectiveness and patient safety).[42] Maxwell proposed extending Donabedian's assessment framework through six 'dimensions' of healthcare quality: access, relevance, effectiveness, equity, acceptability, and efficiency.[43]

Assessment of structures is regarded as the weakest method for assessment of quality of healthcare as there is not typically any established link between structures and clinical outcomes.[44] Assessment of outcomes is a necessary part of monitoring health services but is not always clearly linked to the quality of care delivered.[44] Adverse outcomes would not be expected to arise each time the delivery of care was not of a required standard.[45,46] Equally, adverse outcomes may appear to be more frequent in settings even where care is of high-quality. For example, a surgeon who treats patients at high risk of complications may appear to have worse outcomes than another surgeon who treats only low-risk cases. Adjustment for this case mix is usually complex and may be only partially effective due for example to differences in measurement of variables used in the adjustment process.[47] Mant argued that, for successful implementation of outcome indicators, improved (and standardized) data collection was needed, with validated case mix adjustment methods. He argued that this was only appropriate in situations where there was likely to be sufficient variation in healthcare provision that there might be clinically important variation in outcomes that occurred with sufficient frequency to be identifiable.[47] In general, use of process measures to identify and monitor quality of care has been preferred due to various factors including (i) reduced

complexity compared to outcome measures, (ii) measurable from routine sources; (iii) speed of implementation, and (iv) relevance to healthcare providers.[48] Conversely they can be hard for patients to interpret or identify with, and are not easily aggregated into a useful summary measure.[48] Process measures are only appropriate and relevant where there is high-quality evidence that links the process to improved patient outcomes.[44] For the same reasons as for outcomes, it has been suggested that case mix adjustment also be used in the assessment of care processes such as prescribing[49] and referral.[50,51]

Quality improvement cycles

In general, quality improvement projects require repeated measurements to determine changes in quality over time. This is part of the methodology used in health care quality improvement strategies such as the Plan-Do-Study-Act (PDSA) cycle,[52] which evolved from the original full-cycle medical audit approach.[53] Identification of a need for quality improvement, usually based on some measurement of at least one of the quality domains, would be followed by an improvement intervention and evaluation of its level of success with further measurement of the quality domain(s).

Quality indicators

Some aspects of quality may be assessed through measurement of indicators, an indicator having being 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" (Lawrence and Olesen 1997 p.104[54]).*

There is a sizeable body of research on the development and implementation of quality indicators. Campbell et al. identified in 2002 that, although many quality indicators had been developed for use in hospitals, there was an increasing use in primary care.[55] Indicators may be based upon clinical practice guidelines, or derived *de novo* from primary or synthesised

evidence. Indicators themselves should reflect a potentially measurable aspect of practice – such as structures, processes of care, or outcomes (which may be a true clinical endpoint, an intermediate or surrogate clinical endpoint, or patient-reported other outcome such as satisfaction).[40,43,55-58] The development of clinical quality indicators typically rests on a foundation of evidence relevant to clinical practice, sometimes with the use of consensus exercises to determine the relative importance, validity, and feasibility of candidate indicators.[55,59,60] Although some authorities specify the need for feasibility in data collection for measurement of the indicators,[61] this is not a universal feature of indicator development.

Much work has been done on indicators of quality of care based on the domains of structure, process, and, to an extent, outcome. Studies developing or implementing quality indicators for OA are described fully in Chapter Two.

1.4.3 Primary care services delivery of care for OA

There is some evidence that GPs do not prioritise OA in the context of a consultation, making assumptions that patients do not consider it a priority.[62] A narrative review of comparison of patient experience with GP beliefs about, and attitudes to, OA identified that some GPs do not regard OA as a disease but rather an inevitable consequence of aging,[63] despite the fact that the NICE guidelines state otherwise.[5] Although no evidence was found that an association between GP beliefs about or attitudes towards OA and the diagnosis or care processes received by patients has been identified, it seems plausible that such an association may exist; this would be consistent with behavioural theory such as the theory of planned behaviour.[64,65]

In general, it is well-established that the recorded quality of primary care for OA has been shown to have deficiencies. Porcheret et al. identified through a population survey that recommended core non-pharmacological interventions (exercise, weight loss, written information provision) were used by a minority of survey participants and often on the advice of a friend or relative rather than a health professional.[66] Quality of care for OA has also been

described through the use of studies that included a review of the narrative component of medical records (whether electronic or paper-based), such as by Kirk et al.,[67] Steel et al.,[68] and Broadbent et al.[69] Although this approach is very thorough and captures all the recorded quality of care, it is cumbersome compared to assessments based on electronic retrieval of specific codes and therefore less likely to be used in quality improvement cycles.[53] Previous studies have tended to use a tightly-defined denominator for quality assessment (formally recorded diagnoses of OA) rather than a wider clinical interpretation of persistent joint pain as representing clinical OA.[5] Steel et al. [70] used patient self-report of indicators to measure quality of care and this approach was also used in the population survey element of the MOSAICS study.[1] No studies examining quality of care for OA solely through use of the electronic, coded health record were identified.

Broadbent et al. reported achievement of assessment indicators of between 27% (pain) and 43% (function), information achievement rates of 17% (NSAID risks) and 30% (education), and treatment provision achievement of 48% (paracetamol), 59% (oral NSAIDs), and 90% (referral to orthopædics for arthroplasty consideration where conservative measures had failed).[69] Steel et al. used patient self-report of quality indicators to identify adjusted (for weights and clustering) achievement rates of 24.8% (exercise), 17.7% (education), 41.1% (paracetamol), and 35.8% (orthopædic referral); 77.8% of eligible respondents reported that a clinician had discussed the purpose of treatment in arthritis.[70] Multiple studies in the USA have demonstrated suboptimal quality of care for OA.[71-79] Li et al. also identified suboptimal non-pharmacological care for OA in Canada.[80] In Australia, Runciman et al. reported that 43% of clinical encounters delivered appropriate care for OA.[81] Østerås et al. used self-reported measures in a Norwegian survey to identify a median 27% pass rate for individual indicators (interquartile range 12-50%).[82]

Where guideline adherence has been examined, over-use of radiological investigation has been identified,[83] with under-use of non-pharmacological management approaches.[83-85] One study identified better self-reported recommended use of non-pharmacological therapy by

clinicians, in up to 76% of patients (weight reduction).[86] Clinician self-reported adherence to the EULAR knee OA management recommendations was identified as 74.8% for non-pharmacological management and 73.6% for pharmacological, but only 54.2% for the two approaches combined.[87] Variation in practice between clinicians has been identified, with clinicians who have been in practice for less than 20 years less likely to request tests and more likely to recommend exercise and prescribe oral NSAIDs than those in practice for 20 years or over.[88]

Overall, the quality of care for OA as assessed within the UK and internationally has demonstrated significant shortcomings. Although many of the recommended interventions for OA have only small to medium benefits,[4] given the prevalence of OA, the cumulative population effects of deficits in OA management seem likely to be substantial. Previous studies have used either analysis of the whole medical record or patient self-report; this study adds to those through examination of coded information, which is more feasible to use for continuous quality improvement through audit and feedback. Additionally, this study adds to the relatively small amount of evidence examining associations between patient and clinician factors and quality of care for OA.

1.5 Improving quality of care for OA in general practice

Given that there are established, evidence-based, national and international guidelines for the management of OA, there is potential for general practice, and primary care more generally, to deliver high quality care to people with OA. Although there is considerable concern about the application of multiple single-disease guidelines to patients who have multiple morbidities, the core interventions for OA have considerable overlap with recommendations for other long term conditions, such as weight management and exercise for vascular disease and diabetes, and exercise for chronic obstructive pulmonary disease. The principles of the core interventions for OA management should therefore be very familiar to clinicians, as well as to many patients with relevant comorbidities, and given the capacity

to benefit from such interventions across a range of conditions, one might expect that their implementation should not be unduly onerous. In this section, the possible approaches to quality improvement for OA are outlined.

The Centre for Reviews and Dissemination concluded in 1999 that improvement strategies employing multiple approaches were more likely to be effective than those using only one, and that most interventions were effective under some circumstances but none under all.[89] Educational outreach was found to be generally effective in North America but further work was needed to determine its effectiveness in the UK. Reminder systems were identified to be "generally effective for a range of behaviours" (p. 7). Audit and feedback, use of opinion leaders, and other interventions were identified as having mixed effects with a recommendation that they be used "selectively". Reminder systems are discussed in more detail in the development of the enhanced recording template in Chapter Three, section 3.3.

In terms of the implementation of complex interventions in primary care, a recent (2015) systematic review of reviews concluded that despite an extensive body of literature, it remains unclear which implementation strategies are most effective.[90] Features that seemed to be associated with more successful implementation included use of printed educational materials; educational strategies; educational outreach, audit and feedback; practice facilitation; financial incentives; and multidisciplinary opinion leaders.

The optimum method for improving the primary care of OA specifically has not been established. Brand et al. concluded, in a systematic review, that reported effectiveness of complex OA management interventions varied but there was some evidence to support the use of primary care collaborative care models or multidisciplinary case management; where a positive impact had been observed, this tended to be small to moderate.[91] The review identified a lack of information systems to assist OA disease identification and monitoring within populations. The studies reported in this thesis aim to further the development of systems to facilitate this identification and monitoring of OA in general practice populations.

1.6 Thesis aims and objectives

1.6.1 Aim

The aim of the thesis is to identify quality indicators for the primary care of OA and assess the feasibility of their measurement within primary care in England.

1.6.2 Objectives

(i) To review the existing evidence for quality indicators for OA, applicable to general practice (Chapter Two)

(ii) To identify how the quality indicators may be assessed through information routinely recorded in general practice medical records and develop a mechanism for enhanced recording where necessary (Chapter Three)

(iii) To describe patterns of clinical OA in primary care and the possible extent of misclassification of a peripheral joint pain record as OA within a definition of clinical OA (Chapter Four)

(iv) To describe the routinely-recorded quality of care for clinical OA in general practice (Chapter Five)

(v) To describe the quality of care for OA as captured by the enhanced recording mechanism (Chapter Six)

(vi) To estimate the effect of the enhanced recording mechanism on routinely recorded quality of care (Chapter Seven)

(vii) To investigate the feasibility of use of quality indicators in assessment of a cluster-randomised controlled clinical trial of a complex intervention to improve care for OA (Chapter Eight)

1.7 Outline of thesis

The aim and objectives of the thesis were addressed through research undertaken as part of the MOSAICS studies research programme. MOSAICS was a programme composed of various studies, chief amongst which was a cluster-randomised controlled clinical trial of a complex

intervention to improve uptake of NICE-recommended core treatments for OA. The research undertaken for, and reported in, this thesis formed a part of the MOSAICS programme. Although the candidate was involved in the design of the MOSACIS trial, as reported in the full study protocol,[1] the purpose of this thesis is not primarily to report on the design, conduct, or outcomes of the trial. As the MOSAICS trial is reported in a style consistent with the CONSORT trial reporting recommendations[92] in Chapter Eight, only a brief outline is presented here, to provide orientation to the study as referred to throughout the thesis.

This thesis uses the same definition of clinical OA as that in the MOSAICS study protocol.[1] Peripheral joint sites (hip, knee, and small joints of the limbs) for OA were chosen as these are the most common sites for OA to be manifest.[5] Therefore, OA at the shoulder, elbow, and axial skeleton were excluded from this study. Shoulder pain may represent a wide range of alternative conditions including referred pain or soft tissue disease;[93-95] the elbow is also commonly affected by other soft tissue conditions or may be involved in a regional pain syndrome.[95] Although no studies investigating the positive predictive value (PPV) of pain at these sites as a marker of OA were identified, the PPV would be likely to be lowered by the baseline prevalence of OA at these sites as well as the wide range of differential diagnoses. Spinal OA was also excluded as the management of axial pain has developed into an academic and clinical specialism of its own, with back pain management supported by a separate NICE guideline.[96]

MOSAICS was designed to assess the utility of a model OA consultation (MOAC), which was a complex intervention to increase the adherence to the NICE OA management guidance and in particular uptake of its core management interventions. The complex intervention consisted of clinician training (GPs, nurses, and allied health professionals), additional resources for nurse follow-up clinic appointments, and a patient OA guidebook; practices were paid for participation to reflect additional service costs rather than as an incentive to promote guideline adherence.

Practices were selected for the MOSAICS study on the basis of willingness to participate. To be eligible to participate, practices had to (i) use the EMIS clinical computer system, (ii) have at least two general practice nurses (to allow for cross-cover in service provision during periods of leave), (iii) have been willing to have the OA consultation recording template installed, and (iv) willing to undertake the clinician training. The CRN research facilitation staff approached 10 practices to seek participation; eight consented. Reasons given by the two non-consenting practices were (i) involvement in another research study and (ii) recent commitment to teaching medical students.

Table 1-1: MOSAICS practice characteristics

Trial arm	Practice ID	Practice population	% female	% aged ≥ 45 years	IMD 2010 score	IMD 2010 rank
Intervention	1	23,868	50.7	47.4	3.4	31,573
	2	5,810	47.7	35.9	58.9	921
	7	7,206	50.5	48.2	22.4	12,310
	8	4,077	50.3	48.2	11.7	21,868
Control	3	8,170	51.0	53.7	13.2	20,182
	4	8,461	51.4	58.0	3.2	31,702
	5	7,324	50.7	47.7	28.1	9,085
	6	3,978	47.8	52.8	45.7	3,067
Overall		Mean 8,612	50.4	49.0	-	-
England [Source: ONS[97]]		6,891 (estimated 2012 average[98])	50.2[97]	58.2[97]	-	Range 1 = most deprived to 32482 = least deprived

The eight practices who participated in MOSAICS together form a representative sample of general practice in England. There was a spread of practice size, though with one very large practice that was three times the English mean practice size. The sex structure of the practices was broadly reflective of that of England as a whole. The age structure for the practices was generally younger than that of England as a whole, which would be expected to reduce the number of people consulting with clinical OA. There was also a spread of

deprivation, as estimated by the deprivation score and rank for the practice location (individual patient-level deprivation information was not available). None of the eight MOSAICS practices were included in the CiPCA database at any point.

Before the trial period to test the intervention, there were two preceding phases – one to establish a baseline level of care recorded during the six months immediately prior to cluster randomisation of the practices to the trial, and a further phase for 12 months before that to assess recorded care that was naturally-occurring before any meaningful study engagement with the practices.

The primary aim of the MOSAICS trial was to "*determine the clinical and cost effectiveness of the MOAC intervention in patients with OA."* [protocol, p.3[1]]

The secondary study aims were to:

1. "*Describe the* [patient self-reported] *uptake of core NICE OA recommendations* [the assessment of a person with OA and the central management strategies of education, exercise, weight-loss support where appropriate, and first-line analgesic medication from the NICE guidance management diagram (NICE 2008, p8[5])] *in participants aged 45 years and over with joint pain*"

2. "*Test the feasibility of deriving 'quality markers' of OA management* [from electronic health records] *using a new consultation template and medical record review*"

3. "*Develop and evaluate a training package for management of OA by general practitioners (GPs) and practice nurses*"

4. "*Investigate the impact, feasibility and acceptability of the MOAC intervention*."

This thesis addresses the second of these subsidiary aims. Quality markers (indicators of quality of care) were first identified by the candidate through a systematic review of quality indicators, with a narrative synthesis of indicators considered feasible for use in primary care. Identification of the capability of the routine electronic health record (EHR) to measure the

identified indicators was conducted in a local routine database of primary care consultations (CiPCA). As a result, an enhanced consultation recording mechanism (the template) and its trigger codes for OA and peripheral joint pain likely to represent OA in older adults (which formed the Read-code definition of clinical OA) was designed, developed and implemented by the candidate.

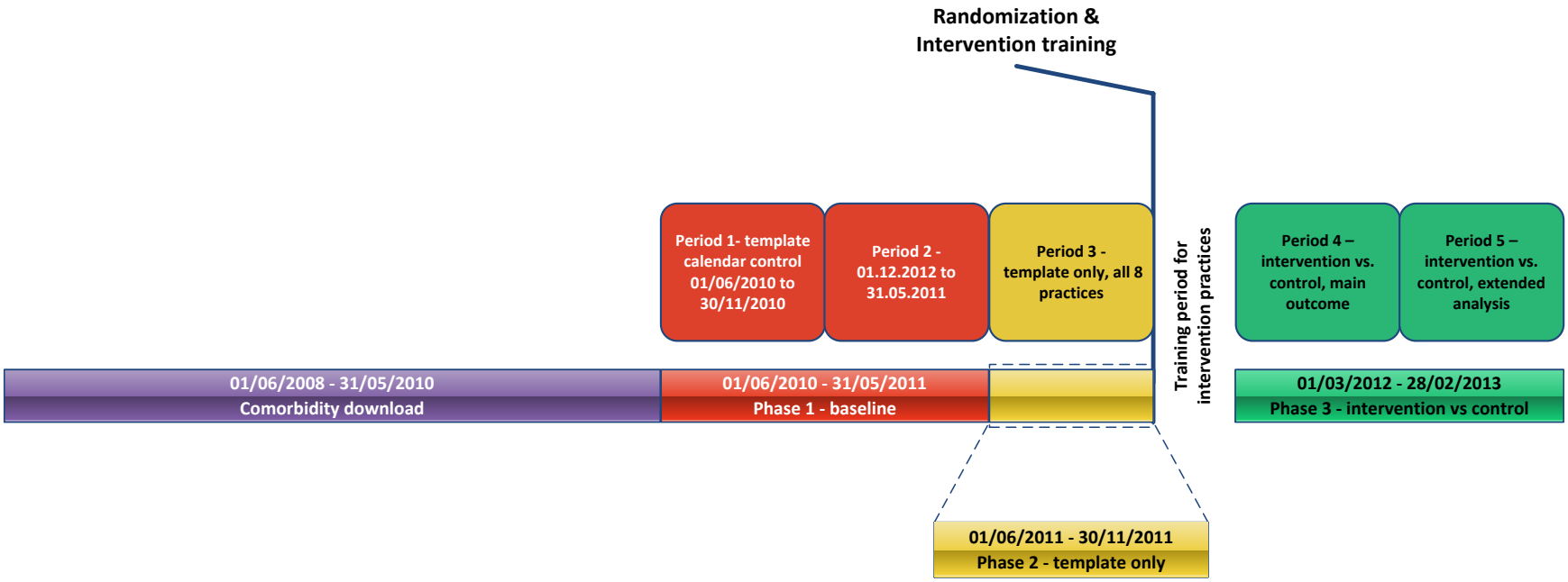
The assessment of baseline coding of clinical OA (defining the denominator population for the routinely-recorded quality indicators) was developed and implemented by the candidate, with an estimate of diagnostic misclassification. The template was installed in all the MOSAICS practices for a six-month period before the clinical trial randomisation. The quality of care for OA was estimated by the candidate both before and after template installation. The estimate of the effect of the template on routinely recorded quality of care was then undertaken by the candidate based upon an analysis by Professor Jordan for the associated publication[99] as shown in Appendix B (for which the candidate was first author, leading the study design and interpretation), though using an extended set of covariates compared to that used in the associated publication.

The feasibility of quality indicators as trial outcomes was assessed through their use as secondary outcome measures in the MOSAICS cluster-randomised controlled clinical trial (during which all practices continued to use the template), including an estimate of the effect of the model consultation. This used a similar approach to the analysis undertaken for the associated publication (Jordan KP et al., in submission), co-designed by the candidate with Professor Jordan; once again the candidate used a broader set of outcomes and covariates for the analysis reported in this thesis, as well as undertaking sensitivity analyses, which extended the analysis submitted for publication. The primary purpose of the associated publication was to assess the effect of the intervention on quality of care rather than the objective here of assessing feasibility of using quality indicators as outcomes.

The MOSAICS study timeline is shown in Figure 1-1.

The discussion (Chapter Nine) summarises the main findings of the thesis, linked to the objectives, and draws conclusions from the work overall. The next Chapter describes a systematic review and narrative synthesis of existing quality indicators relevant to the primary care of OA.

Figure 1-1: MOSAICS study timeline



Chapter Two: Systematic review of quality indicators for osteoarthritis

2.1 Introduction

As discussed in Chapter One, osteoarthritis (OA) is a condition that is not consistently well-managed in primary care. As it is such a frequently-occurring condition with important consequences for patients, health services, and funders, the improvement of condition management for OA is important, and primary care, as the predominant provider of non-surgical care for OA in England, is best-placed to deliver improvements.

Baker highlighted that primary care in general had a shortage of "accessible, valid, complete, and relevant" data in 2000 (p. 83).[100] Although there has been some improvement in data capture in general practice since the advent of computerisation (from the 1990s) and, more recently, the Information Management and Technology (IM&T) Directed Enhanced Service (DES),[101] there is no current nationally mandated structured method to record data regarding OA care. Consequently, quality of care for OA cannot routinely be monitored from general practice data due to a lack of routinely recorded and coded information on the structures, processes of care, and outcomes for OA. Although OA is not unique in this respect, many other long term conditions such as asthma, diabetes, and ischaemic heart disease have been included in the Quality and Outcomes Framework (QOF)[102] and consequently have various general practice processes-of-care or outcomes measurements associated with them. The inclusion of such conditions in the QOF necessitated the development of QOF indicators (essentially quality indicators, covering the four domains of clinical standards, organisational standards, patient experience, and "additional services" agreed-upon by NHS Employers and the BMA[103]). The indicators were then tested with the development of, sometimes complex, business rules to measure the indicators by (see for example the osteoporosis: secondary prevention of fragility fractures ruleset[104]). These business rules have often required the adoption of new coding practices within primary care to ensure that the numerator (people whose care achieves the indicator) and denominator (those eligible to achieve the indicator) can be

correctly identified.[59] New coding practices therefore have an established track record in the implementation of quality measures in general practice in England.

The systematic review reported in this chapter aimed to identify all published studies of the development or implementation of quality indicators for OA care relevant to primary care and to undertake a narrative synthesis to identify the indicators most appropriate for use in general practice in England. This review adds to two previous reviews on quality indicators for OA, by Hochberg,[105] published in 2007, and Strömbeck et al.,[106] in 2013, by identifying indictors applicable to and feasible to measure in primary care and synthesising the disparate indicators within each quality domain into an overarching single indicator, with proposals for implementation in general practice in England.

The review reported here has been published in a peer-reviewed journal (Edwards et al., 2013[107]), reproduced in Appendix A.

2.2 Methods

The methodology used was designed to align with that set out by the National Institute for Health Research (NIHR) Centre for Reviews and Dissemination at the University of York.[108] The candidate led the review, working with two additional reviewers (a second experienced GP and an academic physiotherapist). The methods developed by the candidate were refined following comments from the Research Information Manager and the systematic review team within the Research Institute for Primary Care and Health Sciences (RIPCHS).

2.2.1 Search strategy

The search strategy was developed for use with the NHS Evidence portal.[109] This provided access to the main bibliographic databases, including CINAHL, EMBASE, HMIC, Medline, and PsychINFO. An initial scoping exercise, using a range of OA terms combined with quality indicator terms, was undertaken and identified published articles used to refine the development of the search strategy through use of their medical subject headings (MeSH terms).[110] The scoping exercise also identified repositories of healthcare quality indicators (the Agency for Healthcare Research and Quality

[AHRQ][111]) which were searched in addition to the bibliographic databases. Included articles had reference lists checked for relevant articles not already included. The search was last updated in August 2013.

Box 2-1: Example search strategy for use in Medline

1.	MEDLINE; (qualit* AND (outcome* OR indicat*)).ti,ab
2.	MEDLINE; exp "OUTCOME AND PROCESS ASSESSMENT (HEALTH CARE)"/
3.	MEDLINE; exp QUALITY ASSURANCE, HEALTH CARE/
4.	MEDLINE; exp QUALITY INDICATORS, HEALTH CARE/
5.	MEDLINE; 1 OR 2 OR 3 OR 4
6.	MEDLINE; exp *OSTEOARTHRITIS/
7.	MEDLINE; 5 AND 6
8.	MEDLINE; *ORTHOPEDICS/
9.	MEDLINE; *ORTHOPEDIC PROCEDURES/
10.	MEDLINE; 8 OR 9
11.	MEDLINE; 7 NOT 10
12.	MEDLINE; 11 [Limit to: Publication Year 2000-2013 and English Language]

An example of the search strategy (for Medline) is shown in Box 2-1. The specific terms used were adjusted for the other databases listed above, to account for the differences in controlled vocabulary between databases.

2.2.2 Inclusion and exclusion criteria

Inclusion criteria:

All articles had to refer to patients with OA.

There had to be a focus on either the development or implementation of quality indicators relevant to OA. Such indicators may have been developed through a range of methodologies, but for the

purpose of inclusion, any articles self-identifying or appearing to the reviewers as having undertaken quality indicator development or implementation were eligible.

Articles had to be applicable to a general practice context in England, further described in section 2.2.5. This last criterion was necessarily subjective and was determined by either of the GP reviewers on the basis of professional judgment.

Exclusion criteria:

Articles were excluded if they fell into one of the following categories:

- Any articles not dealing with human subjects
- Any non-English language articles (as no translation facilities were available for this project)
- Articles published prior to 2000 (as the majority of OA management guidance has been published since that time and quality indicators were likely to have been based on treatment guidelines)
- Unpublished articles ('grey literature') with the exception of information from identified quality indicator repositories
- Articles describing the testing of treatment effectiveness and efficacy
- Radiological and surgical techniques, reports, follow-ups (as these were not considered applicable to primary care)
- Basic science articles (not applicable to primary care)
- Case series reports; economic analyses; cost-effectiveness articles; assessment scales; scoring tool development and validation; letters; commentaries; professional development articles; trial protocols (as these were not considered to reflect quality indicator development and implementation)
- Complementary medicine, with the exception of glucosamine and chondroitin, and acupuncture.

2.2.3 Selection of articles

All articles identified through the search strategy above were downloaded to a citation management database (RefWorks[112]). The deduplication facility within this database was used to identify and remove duplicate articles. All titles after deduplication were subject to an initial screening by a single reviewer (the candidate). Articles deemed to be clearly not relevant or eligible according to the inclusion and exclusion criteria were removed from further consideration.

The abstracts of the remaining articles (or, where abstracts were not available, the full-text article) were assessed by the two GP reviewers, including the candidate. Where both reviewers agreed after independent assessment that an article was not relevant or eligible, it was excluded from further consideration.

The final determination of inclusion for articles used full-text information. Each was subject to dual review (by the candidate and one other reviewer). In the case of disagreement, a third opinion was sought from the remaining reviewer. Articles remaining at the end of this process were included in the quality assessment and narrative synthesis phases.

2.2.4 Data extraction

In order to assess the articles which developed or implemented quality indicators, a proforma for data extraction was created and refined after testing by all three reviewers. This structured and standardised the data extraction process, and gathered information relevant to the quality assessment process, including the hierarchy of evidence[113] and consensus method[114] (described in section 2.2.5). Due to differences in the assessment process between development and implementation articles, a separate extraction proforma was developed for each. To improve reliability, each included article was subject to data extraction by two reviewers. Where differences in extraction occurred, these were resolved by consensus and reference back to the original article. Any remaining disagreement was resolved by discussion with the third reviewer, with a third data extraction to be undertaken by the third reviewer if required. The data extraction proformas are shown in Appendix D.

Where information was considered unclear or insufficient, the corresponding authors of the included papers were contacted if necessary for clarification.

2.2.5 Quality assessment process

Although assessment of systematic reviews, randomised controlled trials and ecological studies has developed into a well-defined science, there was no similarly well-established method identified for assessment of quality indicator development and implementation. The assessment process used for this review was based on the Outcome Measures in Rheumatology (OMERACT) filter of truth, discrimination, and feasibility.[115] This tool was originally used to assess clinical trials in rheumatoid disease but its structure was transferrable to other studies, including for example use in the development of quality indicators, guidelines and outcomes in scleroderma.[116]

The quality assessment process occurred at two levels – (i) that of the whole study (or group of studies, where a number of studies were based on the same foundation work), to provide information about the evidence base used to identify the indicator content, the consensus exercise to develop the indicators, target population, method of measurement, implementation, and assessment of reliability; and (ii) that of the individual indicator (to determine its feasibility, for example).

Sometimes, studies did not report the whole development of indicators but rather referred to other articles. Here, the evidence for the group of studies was considered as a whole corpus rather than for each report separately; where derivative indicator sets were developed, any supplementary evidence or consensus building was considered and reported within the review.

The assessment tool was based on a methodology for assessment of clinical performance measures by Geraedts et al.[117] This was the most suitable instrument identified from a literature search for assessment methodologies appropriate to quality indicator development and was also broadly consistent with the COSMIN checklist[118] for assessment studies of methodological quality of health status measurement instruments. Additional literature supporting the assessment methodology is

cited below for each assessment domain. The extraction tools used an approach consistent with the OMERACT filter, with various measures for each OMERACT domain.

Assessment at the level of the study or study group

Articles, or groups of articles if produced by the same development group, were assessed against the following criteria:

- Conflict of interest in the development or implementation processes was assessed from any statements of conflict of interest in the article; if no statement was made, the reviewers' judgment was used to assess whether it was likely any conflict of interest would have affected the study. [OMERACT domain: truth]
- Current relevance: In order to be considered 'current', included indicators should have been developed since the NICE OA guidance publication in 2008 or, if developed earlier, the reviewers had to determine whether it was still relevant to best practice. Evidence of a mechanism for keeping the indicator set up-to-date with new evidence was sought. [OMERACT domain: truth]
- Content validity (i): Method of evidence collection and synthesis for the indicator development (where a systematic review was best and expert opinion was least robust, as set by the University of Oxford Centre for Evidence-based Medicine[113]). This was especially relevant to interventions offered to patients such as drugs or devices, where robust evidence of potential benefits and risks was necessary for good clinical care. [OMERACT domain: truth]
- Content validity (ii): Consensus was defined in the first Assessing Care Of Vulnerable Elders (ACOVE-1) indicator set[119] as "adequate scientific evidence or professional consensus supported a link between the process specified by the indicator and a health benefit to the patient" (p.649). The method of consensus development of the indicators was identified (the modified RAND process was regarded as the optimum method, followed by the Delphi method[114]). [OMERACT domain: truth]

- Field testing (feasibility): Articles were considered to offer a higher standard of indicator if there was evidence that the indicators had been tested or implemented in practice. [OMERACT domain: feasibility]
- Test-retest reliability: this was assessed through implementation studies to see whether multiple data extractions of quality indicator results from the same dataset gave the same results. [OMERACT domain: truth]

Assessment of individual indicators

- External validity: (defined by Geraedts et al.[117] for clinical performance measures as "*The measurement activity and subsequent orientation of the care process measured towards established performance thresholds actually should lead to an improvement of medical care delivery and/or of the outcomes of care*" [p.81]) was assessed by examination of implementation studies for evidence of quality improvement in populations in which the indicator had been used.[117] [OMERACT domain: truth]
- Responsiveness: (the identification of quality improvement over time in a population independently known to have sustained such an improvement) was assessed through evidence from implementation studies. [OMERACT domain: discrimination]
- Feasibility: (including identification of the applicable population and method of measurement) was assessed through any recommendation regarding the method for implementation in the development studies, or through implementation studies, and whether such methods were judged compatible with routine general practice in England (which required that the indicator be consistent with the NICE OA management guidelines[5]). This domain was deemed to include aspects of practice such as the clinical or administrative burden, additional cost, and likely acceptability to clinicians and patients. This component was judged by either of the GP raters through clinical practice expertise rather than by comparison with identified empirical evidence, though factors such as methods of prior implementation and data sources were important components in this decision (data needed in principle to be retrievable from primary care clinical information

systems, with or without some degree of modification to current practice). [OMERACT domain: feasibility]

- Reproducibility: assessed by identification of similar indicators independently created from more than one study. [OMERACT domain: truth]

2.2.6 Narrative synthesis

Identified indicators were collated in themes derived from the NICE OA guidance[5] such as patient assessment, non-pharmacological management, pharmacological management, and specialist assessment. All indicators were arrayed within themes, divided into subthemes by their specific focus (such as assessment of pain or assessment of function).

On the basis of this array, an exemplar indicator was selected from those identified, determined by the robustness of its development, implementation, and applicability to general practice in England. Differences of opinion between the reviewers was resolved by consensus.

Each subtheme was synthesised into an indicator which, in the reviewers' opinion, was suitable for direct implementation. This entailed the clear articulation of a quality criterion, numerator, and denominator, as is customary for quality indicator development.[120] This final indicator was based upon the exemplar but was also intended to overcome any heterogeneity between the published indicators that were deemed valid for inclusion (such as timeframes for assessment), be consistent with the NICE guidance,[5] and be suitable for implementation with only modest modification to usual practice (i.e. feasible).

2.3 Results

The stages of analysis are shown in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart[121] From the original search strategy after deduplication, 10,853 articles were identified. (Figure 2-1). The final inclusion set numbered 32,[67-80,82,111,119,122-136] with 14 development articles and 18 implementation articles resulting in 10 indicator groups. Two corresponding authors were contacted for clarification of issues, one of whom responded.

2.3.1 Quality indicators: assessment

The 32 included articles were put into ten groups, allocated on the basis of their original development or modification. The flow of information from studies to indicators is illustrated in Figure 2-2 and the ten indicator development groups are shown in Table 2-1 as well as Table 2 of the associated publication [Appendix A]. Many aspects of the quality of the indicators themselves were assessed at group level: method of evidence synthesis; consensus methodology used; target population; method of measurement of the indicator; any testing or implementation of the indicators. Reliability assessments for data extraction in implementation studies, where reported, were given at the level of the study as a whole rather than the individual indicator. Some elements of feasibility assessment were also reported at this level, such as any methods used to test or implement the indicators.

Figure 2-1: PRISMA reporting flowchart for stages of the systematic review

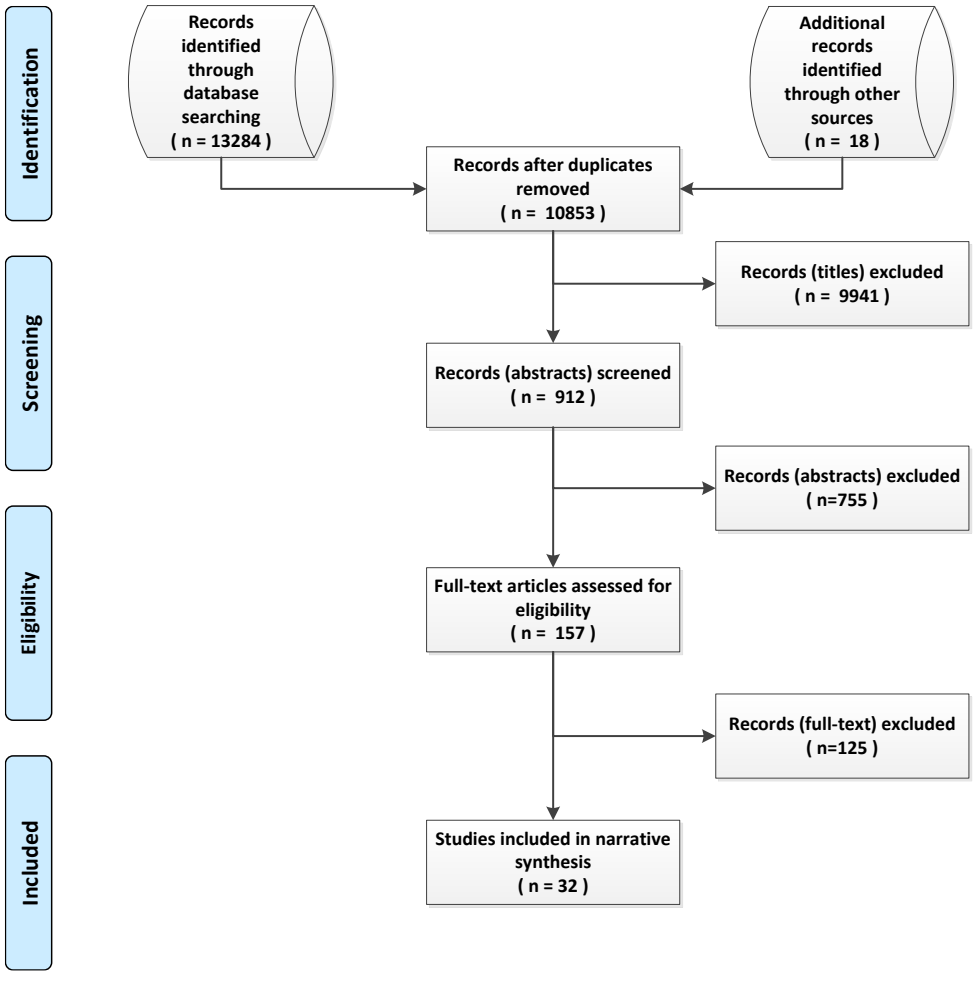
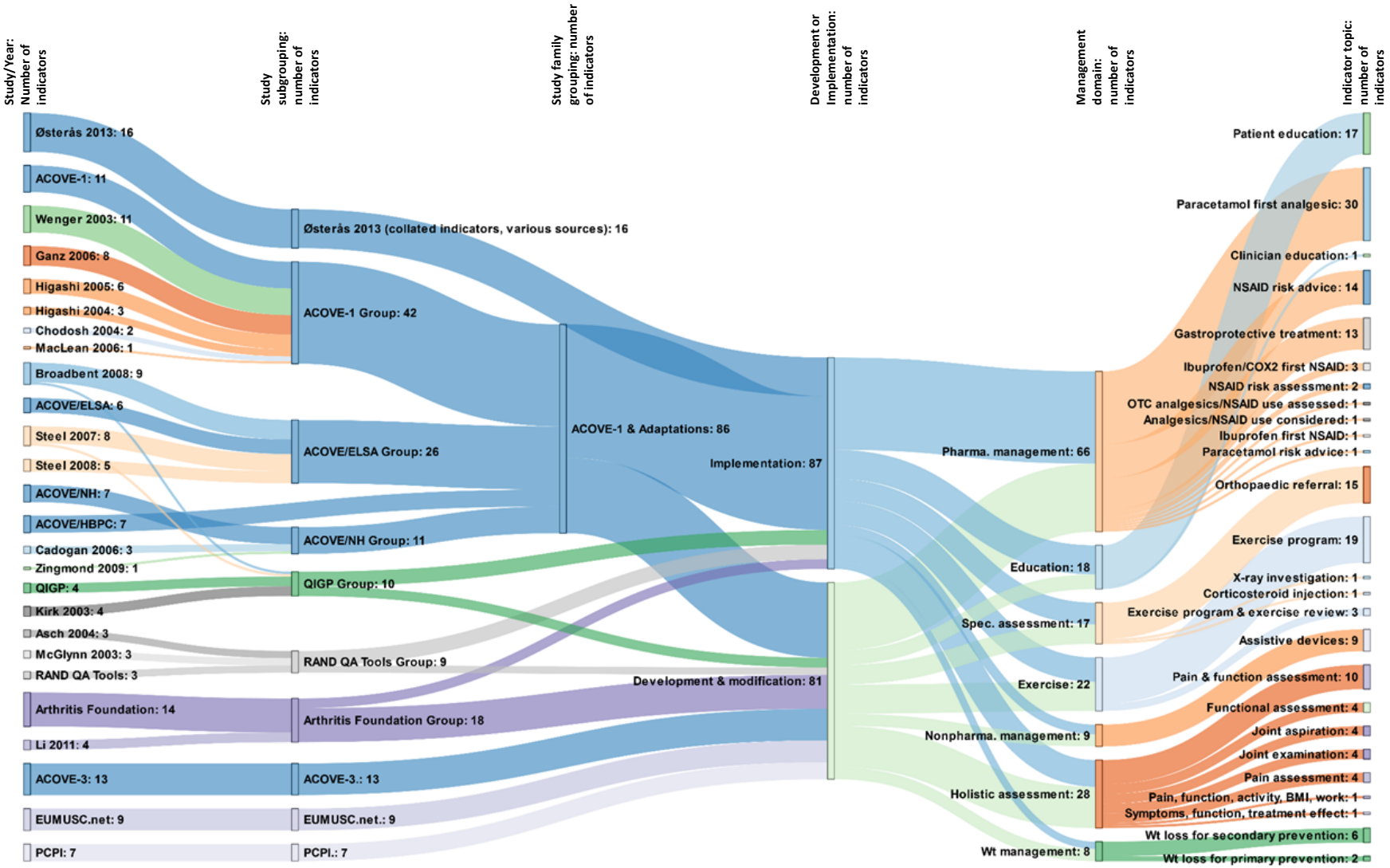


Figure 2-2: Sankey diagram of flows from included studies to synthesised indicators



Conclusions on some aspects of quality assessment were common across all indicators and so are not separately reported. These were:

- Conflict of interest: Although an absence of conflict of interest was not explicitly stated for every article, the reviewers considered that there was no significant likelihood of resulting bias in the results
- Current relevance: no studies identified a method of ensuring updates of indicators in the light of new evidence, though the ACOVE-1 indicators were subsequently updated in ACOVE-3
- External validity and responsiveness: these aspects had not been demonstrated in any of the indicator development or implementation studies

No articles were excluded from the final set as the result of the quality assessment process.

Table 2-1: Quality indicators grouped by developer, with quality assessment

Indicator/group development Author & Date	Truth		Target population	Method of measurement	Feasibility	
	Content validity (i): Evidence synthesis	Content validity (ii): Consensus method			Test-retest reliability	Indicator testing or implementation
1. RAND Quality of Care Assessment Tools (RAND QA) Moore 2000[135]	✓	✓✓	Not specified.	Medical record review.	✓	✓
2. Assessing Care of Vulnerable Elders (ACOVE) -1 MacLean 2001[119,134,]	✓✓	✓✓	Vulnerable elders	Not specified.	✓	✓
3. ACOVE-1 adapted for the English Longitudinal Study of Ageing (ELSA) Steel 2004[124]	✓	✓✓	Older patients (≥65y) in the UK	Interviews for the English Longitudinal Study of Aging.	✓	✓
4. ACOVE-1 adapted for Nursing Home (NH) implementation (ACOVE/NH) Saliba 2005[133]	✓	✓	Long-stay NH residents ≥65y.	Not specified.	✓	✓
5. ACOVE-1 adapted for the Home-based Primary Care Quality Initiative (HPCQI) Smith 2007[123]	✓	✓	Patients ≥60y who are homebound.	Not specified.	x	x
6. ACOVE-3 ACOVE investigators 2007[122,126,132]	✓✓	✓✓	Community-dwelling individuals aged ≥65y who are at greater risk of death or functional decline over a 2-year period.	Medical records and/or administrative data, patient or proxy interview.	✓	✓

(con't)

Indicator/group development Author & Date	Truth		Target population	Method of measurement	Test-retest reliability	Feasibility
	Content validity (i): Evidence synthesis	Content validity (ii): Consensus method				Indicator testing or implementation
7. Quality Indicators for General Practice (QIGP) Underwood 2002[131]	✓	x	Not specified.	Not specified.	✓	✓
8. Arthritis Foundation Arthritis Foundation 2004[127,128]	✓	✓✓	Patients with OA.	Not specified.	x	✓
9. PCPI Physician Consortium for Performance Improvement (PCPI) 2006[129]	x	x	All patients aged ≥21y with a diagnosis of OA.	Medical record data extraction (detailed numerator and denominator information provided)	x	x
10. EUMUSC.net 2012[125]	x	x	All adult patients with OA of hand, hip or knee.	Varies. Examples include patient record or survey. Numerator and denominator clearly identified.	x	x

Key - ✓✓: optimal methodology; ✓: acceptable methodology; x: either inadequate evidence of method used or inadequate methodology (Adapted from Edwards et al. 2015[107]).

Of the ten groups, five were found to be based upon the ACOVE indicator series. These indicators were found most closely to fulfil the indicator quality specifications noted under *2.2.5 Quality assessment process*, above. The strengths of the ACOVE indicators were particularly in their robust evidence collation and synthesis, the consensus methodology used in the indicator development, field testing in implementation studies (ACOVE-1 and 3), and in their update cycle in ACOVE-3. Relatively minor modifications to ACOVE-1 were made for several other study groups: the English Longitudinal Study of Aging,[124] assessment of care quality in nursing homes,[133] and for home-based primary care.[123] Such modifications included those to the target population and to the timeframe for recommended processes of care. Although the degree to which the modifications were the result of further empirical study and consensus was variable, the identified modifications were judged to be compatible with the original parent indicators. Some heterogeneity was noted regarding the use of oral NSAIDs and gastroprotection agents, with respect to the particular drugs recommended or the target population.

The remaining indicator groups were less well matched to the quality assessment criteria. The RAND indicators,[135] the earliest identified, were based upon a literature review but this was not identified as a systematic review, although the consensus exercise was high-quality. The Arthritis Foundation indicators[127,,128] were based upon a "comprehensive" (rather than systematic) literature review and a high-quality consensus process; one implementation of a non-pharmacological Arthritis Foundation indicator (weight loss) was identified. The remaining indicator groups used either a less rigorous evidence base or consensus exercise, or did not specify how the indicators were derived. Some had no evidence of any testing or implementation (HBPCQI[123], PCPI[129] and EUMUSC.net[125]), though the EUMUSC.net indicators had only recently been published at the time of the review.

All identified indicators predominantly used processes of care. The EUMUSC.net[125] indicator set also included three outcome measures, relating to a proposed 20% improvement in pain and function on patient-reported outcome measures within three months of commencing treatment, and workforce participation.

The studies from the full-text assessment that did not meet the inclusion criteria for the narrative synthesis are listed in Appendix D.3.

2.3.2 Quality indicators: narrative synthesis

The indicator themes are listed, with a chosen exemplar indicator, evidence of an indicator's reproducibility in other studies, all identified instances of implementation, and an assessment of feasibility, in Table 2-2. A full breakdown of all identified indicators within holistic assessment is shown in Appendix D.4, to demonstrate the way in which the original indicators were brought together under each unifying proposed indicator for implementation. The themes are discussed further below, with the conclusions from the review process. The final proposed indicators are shown in Table 2-3.

Holistic assessment

There were 28 instances of indicators relating to holistic assessment in 17 studies (nine development and eight implementation studies), as identified in Table 2-2. Indicators relating to assessment of pain and function occurred relatively frequently and were considered valid and feasible. The ACOVE-3 examples were considered most closely to fulfil the assessment criteria (due to the systematic review evidence collection methodology, modified RAND appropriateness consensus method, and field testing with a high level of feasibility for implementation with only small modification to usual practice) and form the exemplar indicators for the pain and function sub-themes. The proposed final indicator wording[107] was *the percentage of patients with a working diagnosis of OA with evidence of pain assessment within the previous 12 months*, and *the percentage of patients with a working diagnosis of OA with evidence of function assessment within the previous 12 months*.

Indicators for joint examination and aspiration were identified, though less frequently. Although these too were the outcome of studies with high-quality evidence synthesis and consensus development, they had not been successfully implemented and were considered less feasible for use in general practice in England.

The indicators for holistic assessment are shown in Appendix D.4, with the thematic extractions and resulting indicators proposed for implementation as an example of the synthesis output.

Education and information

Indicators within this theme occurred 18 times in 10 studies (four development and six implementation studies). The Arthritis Foundation indicator was selected as the exemplar, being based upon a high-quality evidence synthesis and consensus exercise as well as being consistent with the similar ACOVE-1 indicator (but more recent). It was also considered to be consistent with the EUMUSC.net education indicator. The ACOVE-3 OA indicator set did not include an education indicator, though the medication use and pain management sets did include such indicators.[137,138] These medication use and pain management indicator sets from ACOVE-3 did not meet the inclusion criteria for this review as they were not stated specifically to refer to OA but rather were generic medication use and pain management indicators; the five indicators in the medication use set that were simultaneously included in the ACOVE-3 OA indicator set were included in this review.[132]

Within the theme, some variability was noted relating to the recommended timeframe specified for OA education processes to occur within. For example, the ACOVE-1 set specified slightly different indicators for patients with incident OA (education to be offered within six months of diagnosis) and prevalent disease (a record of education should be present for people who have had symptomatic OA for 12 months).[134] The Arthritis Foundation set did not differentiate between incident and prevalent disease, specifying that education should be provided or recommended at least once for people with a diagnosis of symptomatic OA for at least three months.[127] Other indicators did not include a timeframe.[82,124]

Most indicators referred to the need for education about the natural history, treatment, and self-management of the disease. The EUMUSC.net indicator referred to *continuous access to education on important preventive and therapeutic strategies in the management of OA*, and some implementation indicators were less specific: "*Has any doctor or nurse ever talked to you about*

what the specific purpose of the treatment for your arthritis or joint pain is?" (Steel 2008,[70] additional Table A). Østerås[82] (p1046) used five education indicators – "*Have you been given information about how the disease usually develops over time?",* "*Have you been given information about different treatment alternatives?*", "*Have you been given information about how you can live with the disease?*", "*Have you been given information about how you can change your lifestyle?*", and "*Have you been given information about the importance of physical activity and exercise?*" There was some variation in the requirement for education to be "given" or "recommended". The method for collecting the information required to assess indicator achievement was generally not clear. In one implementation study, the patient was interviewed by telephone: "*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?*"[74] The indicator in this study was considered achieved if there was at least one positive response. The reviewers rated this indicator as less feasible for implementation (as it would require either a detailed set of education indicator records or a series of patient self-report indicators). The reviewers proposed a more general education indicator: *the percentage of patients with a working diagnosis of OA with evidence of education or advice since diagnosis*.

The EUMUSC.net indicator set[125] also included an indicator for clinician education. This was not considered eligible for inclusion in the systematic review as it was not patient-focussed.

Exercise and physiotherapy

Twenty-two instances of indicators recommending or prescribing exercise or physiotherapy were identified in 18 studies (nine development and nine implementation). One related to patients with "symptomatic OA," [125] one to patients with hand, hip or knee OA,[82] six to patients with hip or knee OA [71,77,80,128,132,135] and the remainder to patients with knee OA.

There was variation in the forcefulness of the indicator. Some recommended only that exercise be "recommended" or "considered", while others required it to be "prescribed". These terms were

somewhat open to interpretation. When considering feasibility, it seemed likely that the different terms would result in different clinical behaviours were the indicators to be implemented. There would also be different implications for data capture: use of "recommended" or "considered" may risk becoming a 'tick-box' requirement, whilst "prescribed" may require evidence of prescription (possibly written instructions or a referral to an exercise programme or physiotherapist).

The type of exercise referred to by the indicators was also noted to be variable: specific exercise programmes, general aerobic exercise, or referral to a physiotherapist. The criterion for success in one study was a record of prescription for lower extremity strengthening or ambulation with a physical therapist or restorative nursing assistant after OA diagnosis.[72] Others used special data sources such as patient interview and sometimes the assessment method was not specified. From implementation studies, the reviewers concluded that feasible indicators for primary care relate to the offer of exercise advice or referral to a physiotherapist, and review of current exercise activity. It was considered feasible to separate two elements of the ACOVE-3 exemplar indicator into a proposed indicator for advice, recommendation, or prescription of exercise (*the percentage of patients with a working diagnosis of OA in the hip or knee with evidence of exercise advice or physiotherapy referral since diagnosis*), and a proposed indicator of annual review of activity (*the percentage of patients with a working diagnosis of OA with evidence of an activity review within the previous 12 months*).[107]

Weight loss

Weight loss indicators for overweight patients occurred eight times in five studies (three development, two implementation). Six instances related to patients with established OA and two (from the Arthritis Foundation, ACOVE-3) to primary prevention of OA.

Variability was noted in the BMI threshold for intervention ("overweight" not further defined, BMI ≥27 kgm^{-2}, or BMI ≥30 kgm^{-2}). As with exercise and physiotherapy, there was variation in the type of intervention contained within the indicator, from advice only to referral to a formal weight loss programme. Two implementation studies between them implemented weight loss indicators three

times. Li et al.[80] implemented the Arthritis Foundation indicator for weight loss in symptomatic OA, chosen here as the exemplar indicator, by using entry to a weight loss programme or having a dietetics appointment as criteria for success. Østerås et al.[82] used patient self-report of either advice to lose weight or referral for help with weight loss as success criteria.

The reviewers considered that a record of advice to lose weight would be a feasible indicator. An indicator regarding referral to a weight loss programme for patients with OA who had been overweight for three years or more, such as in the second Arthritis Foundation weight loss indicator, would be less feasible to implement due to difficulty in identification of the denominator population. The indicators proposed for implementation are *the percentage of patients with a working diagnosis of OA with a BMI ≥25 kgm^{-2} who have a record of weight loss advice within the previous 12 months* and, for primary prevention of OA, *the percentage of patients with a BMI ≥30 kgm^{-2} who have a record of weight loss advice within the previous 12 months*.[107]

Assistive devices

Indicators for the assessment of need of ambulatory or non-ambulatory assistive devices occurred nine times in five studies (three development, two implementation studies). No interventions resulting from the assessment of need were included in the indicator. The ACOVE-3 indicators were chosen as exemplars; implementation of the Arthritis Foundation indicators (which were consistent with the ACOVE-3 indicators) was considered to provide evidence of feasibility. Li et al.[80] used a physiotherapy or occupational therapy (OT) consultation within the previous year as criteria for success for ambulatory or non-ambulatory devices respectively; Østerås et al.[82] used patient self-report of assessment for assistive devices.

General indicators for the assessment of, or referral to physiotherapy/OT, seem feasible for use in primary care. The proposed implementation wording for the ambulatory and non-ambulatory assistive device indicators is *the percentage of patients with a working diagnosis of OA with evidence of functional impairment who are recorded as receiving a referral or assessment for ambulatory assistive devices within the previous 12 months* and *the percentage of patients with a*

working diagnosis of OA with evidence of functional impairment who are recorded as receiving a referral or assessment for assistive devices within the previous 12 months.[107]

Analgesics

Indicators regarding analgesics were the most frequently occurring, with 53 instances in 22 studies (11 development, 11 implementation). Topics included assessment of current use of analgesics, consideration of analgesics, use of appropriate first-line agents (paracetamol), and risk assessment and communication. Exemplar indicators were generally derived from high-quality evidence synthesis and consensus methods (although the PCPI NSAID risk assessment indicator[129] had a less clear evidence and consensus basis, it was consistent with a similar indicator from ACOVE-1). Variability in the indicators regarding the use of oral NSAIDs use was noted, particularly in terms of the drugs recommended: some did not specify which NSAIDs should be used, others recommended ibuprofen,[131] or ibuprofen or a COX-2 inhibitor.[68,69] There was a similar variation in recommendation about risk assessment and advice.

Four indicators relating to use of paracetamol or oral NSAIDs were considered feasible for implementation:

- *the percentage of patients with a working diagnosis of OA with evidence of paracetamol as the first oral analgesic prescribed or advised since diagnosis*

- *the percentage of patients with a working diagnosis of OA taking oral analgesics or NSAIDs with evidence that a suitable maximal dose of paracetamol was tried beforehand*

- *the percentage of patients with a working diagnosis of OA with evidence of a standard NSAID or COX-2 inhibitor as the first oral NSAID prescribed or advised since diagnosis*

- *the percentage of patients with a working diagnosis of OA taking an oral NSAID with a documented risk assessment prior to first prescription*.

No indicator for the use of topical NSAIDs was identified. Indicators for the assessment of existing use of, and consideration of further treatment with, analgesics,[129] and another implemented

indicator for stronger analgesia[82] were not selected due to an unspecified evidence base and consensus approach, although their face validity was acknowledged by the reviewers. Indicators relating to risk communication were not selected as feasible due to difficulties in implementation in the electronic medical record as these were considered to require free-text record analysis, as they would not otherwise be easily and meaningfully recorded.

Gastroprotection

There were 13 occurrences of gastroprotection indicators in 12 studies (four development, eight implementation). There was substantial variation in the triggers specified for gastroprotection use as well as in the choice of agent to be used. The most general indicator was developed in a study which cited a meta-analysis to support the use of gastroprotection agents in reducing the incidence of adverse events in people taking oral NSAIDs.[129] This was consistent with the conclusion of the NICE guidelines[5] that all patients with OA over the age of 45 years should be co-prescribed a proton pump inhibitor (PPI) when also prescribed an oral NSAID.

Where indicators have been implemented, the denominator (eligible) population was frequently determined by reference to past medical history (e.g. history of peptic ulcer disease) or co-therapy (aspirin, warfarin).

The PCPI indicator[129] was considered the most relevant and feasible ("*Percentage of patient visits for patients aged 21 years and older with a diagnosis of OA during which GI prophylaxis was considered"*), with some changes for the final proposed indicator to reflect the NICE guidelines that a PPI should routinely be used with an oral NSAID for adults with OA: *the percentage of patients with a working diagnosis of OA taking an oral NSAID who are also prescribed a PPI or alternative gastroprotective agent*. The inclusion of alternative gastroprotective agents was made to reflect the clinical reality that PPIs are not always suitable or tolerated.

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

Sixteen indicator occurrences in 14 studies (six development, eight implementation) were identified. These 16 indicators mainly related to consideration of X-ray use and more specialist

interventions (specialist referral, joint injection) where symptoms were not controlled by other means.

As NICE guidance for OA did not recommend routine use of X-rays, none of the identified X-ray indicators were considered relevant or feasible for primary care in England.

A number of indicators referred to the use of failure of conservative treatment as a prerequisite for onward specialist referral, though the term 'failure' was not clearly defined. One implementation study asked patients if they had pain and functional impairment, and had been offered a joint replacement or orthopædic assessment.[74] Another used a patient self-report to identify failure of conservative treatment leading to referral.[82]

An indicator, based on the exemplar ACOVE-3 indicator, mandating that all other indicators must have been recorded as appropriately met prior to referral was considered to be feasible for primary care in England: *the percentage of patients with a record of achievement of all other applicable indicators prior to specialist referral*.

Outcome indicators

The EUMUSC.net indicator set [125] included three outcome measures:

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

Whilst acknowledging the desirability of outcome measures in addition to process-of-care measures, the reviewers considered these to be less feasible for use in general practice in England due to difficulties in accounting for comorbidities and case mix, as well as the fact that the required patient-reported outcome measures were not in established use in general practice.

Table 2-2: Narrative synthesis of exemplar indicators and their feasibility for use in primary care

Overarching theme [guidance source]	Exemplar indicator	Reproducibility (studies also developing similar indicators)	References for identified implementation studies	Feasibility assessment
Holistic Assessment: Pain [EULAR (hand, hip, knee), NICE]	"IF a vulnerable elder 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][122,126,132]	RAND QA[135], ACOVE-1[119,,134], and as adapted (ELSA[124], HPCQI[123]), Arthritis Foundation[127,,128], PCPI[129], EUMUSC.net[125]	Asch[71] Broadbent[69] Chodosh[73] Ganz[74] McGlynn[77] Osteras[82] Steel 2007[68] Wenger[78]	Feasible: Requires change in routine coding to improve capture of this information.
Holistic Assessment: Function [ACR (hand), EULAR (hand, hip, knee), NICE]	"IF a vulnerable elder 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][122,126,132]	RAND QA[135], ACOVE-1[119,,134], and as adapted (ELSA[124], HPCQI[123]), Arthritis Foundation[127,,128], PCPI[129]	Asch[71] Broadbent[69] Chodosh[73] Ganz[74] McGlynn[77] Osteras[82] Steel 2007[68]	Feasible: Requires change in routine coding to improve capture of this information.
Holistic Assessment: Examination [NICE, EULAR (hand, hip, knee)]	"IF a patient is begun on a drug treatment for "joint pain," "arthritis," or "arthralgia," THEN evidence that the affected joint was examined should be documented." [Arthritis Foundation][127,,128]	ACOVE/NH[133] (2 indicators, one relating to new residents, one to residents prescribed a drug to treat new joint pain), PCPI[129]	This indicator had not been implemented in any identified studies	*Not suitable: This indicator was not considered feasible for implementation due to limitations in recording examination findings in format suitable for easy audit– could only be implemented with substantial change to coding behaviour.*

(con't)

Overarching theme [guidance source]	Exemplar indicator	Reproducibility (studies also developing similar indicators)	References for identified implementation studies	Feasibility assessment
Holistic Assessment: Joint aspiration [None]	"IF a vulnerable elder has monoarticular joint pain associated with redness, warmth, or swelling AND the patient also has an oral temperature greater than 38.0°C and does not have a previously established diagnosis of pseudogout or gout, THEN a diagnostic aspiration of the painfully swollen red joint should be performed that day…" [ACOVE-1][119,,134]	ACOVE/NH [133], HPCQI [123]	Although implemented in one identified study[78], no patients were eligible for this process of care (numerator/denominator of zero)	*Not suitable: This indicator was not considered feasible for implementation due to limitations in recording differential diagnoses, and because such patients are likely to be referred to secondary care as an emergency; denominator hard to define.*
Education [EULAR (hand, hip, knee), NICE, OARSI]	"IF a patient has had a diagnosis of symptomatic osteoarthritis of the knee or hip for > 3 months, THEN education about the natural history, treatment, and self-management of osteoarthritis should have been given or recommended at least once…" [Arthritis Foundation][127,,128]	ACOVE-1 (2 variations – new and pre-existing disease)[119,,134], and as adapted (ELSA[124]), EUMUSC.net[125]	Broadbent[69] Ganz[74] Osteras[82] Steel 2007[68] Steel 2008[70] Wenger[78]	Feasible: Requires change in routine coding to improve capture of this information.
Exercise 1 & 2 (ACR (hip, knee), EULAR (hand, hip, knee), NICE, OARSI	"IF an ambulatory vulnerable elder 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][122,126,132]	<u>Initial recommendation</u> RAND QA [135], ACOVE-1 (indicators for new and pre-existing disease)[119,,134], and as adapted (ELSA[124], ACOVE/NH[133], HPCQI[123]), Arthritis Foundation[127,,128], PCPI[129], EUMUSC.net[125] <u>Annual review</u> RAND QA[135], ACOVE-1[119,,134]	<u>Initial recommendation</u> Asch[71] Cadogan[72] Ganz[74] Higashi 2005[136] Li[80] McGlynn[77] Osteras[82] Steel 2008[70] Wenger[78] <u>Annual review</u> Li[80]	Feasible: Requires change in routine coding to improve capture of this information.

(con't)

Overarching theme [guidance source]	Exemplar indicator	Reproducibility (studies also developing similar indicators)	References for identified implementation studies	Feasibility assessment
Weight loss 1 [ACR (hip, knee), NICE, OARSI	"IF a vulnerable elder is obese (body mass index (BMI) ≥30 kgm^{-2}), THEN he or she should be advised annually to lose weight…" [ACOVE-3][122,126,132]	Arthritis Foundation[127,,128]	No implementation studies identified for this indicator.	Feasible: Should be captured from existing weight and health promotion records.
Weight loss 2 [ACR (hip, knee), NICE, OARSI	"IF a patient has symptomatic osteoarthritis of the knee or hip and is overweight (as defined by body mass index of ≥27 kgm^{-2}), THEN the patient should be advised to lose weight at least annually AND the benefit of weight loss on the symptoms of osteoarthritis should be explained to the patient…" [Arthritis Foundation][127,,128]	EUMUSC.net[125]	Li[80] Osteras[82]	Feasible: Consider a lower BMI threshold of 25 kgm^{-2} for consistency with the usual definition of 'overweight'. Should be captured from existing weight and health promotion records.
Weight loss 3 [ACR (hip, knee), NICE, OARSI]	"IF a patient has symptomatic osteoarthritis of the knee or hip and has been overweight (as defined by body mass index of ≥27 kgm^{-2}) for 3 years, THEN the patient should receive referral to a weight loss program…" [Arthritis Foundation][127,,128]	No other variations	Osteras[82]	*Not suitable: This indicator was considered less feasible for implementation in primary care using routine data sources.*
Aids and devices 1 [ACR (hip, knee), EULAR (hip, knee), NICE, OARSI]	"IF a vulnerable elder has symptomatic OA of the hip or knee and has difficulty walking that makes activities of daily living difficult for longer than 3 months, THEN the need for ambulatory assistive devices should be assessed…"[ACOVE-3][122,126,132]	Arthritis Foundation[127,,128], EUMUSC.net[125]	Li[80] Osteras[82]	Feasible Requires change in routine coding to improve capture of this information.
Aids and devices 2 [ACR (hand), NICE]	"IF a vulnerable elder has symptomatic OA and has difficulty with non-ambulatory activities of daily living (ADL), THEN the need for ADL assistive devices should be assessed…"[ACOVE-3][122,126,132]	Arthritis Foundation[127,,128], EUMUSC.net[125]	Li[80] Osteras[82]	Feasible Requires change in routine coding to improve capture of this information.

(con't)

Overarching theme [guidance source]	Exemplar indicator	Reproducibility (studies also developing similar indicators)	References for identified implementation studies	Feasibility assessment
Paracetamol 1 [ACR (hip, knee), EULAR (hand, hip, knee), NICE, OARSI]	"IF a vulnerable elder is started on pharmacological therapy to treat OA, THEN acetaminophen should be tried first…"[ACOVE-3][122,126,132]	RAND QA[135], ACOVE-1[119,,134], and as adapted (ELSA[124], ACOVE/NH[133], HPCQI[123]), QIGP[131], Arthritis Foundation[127,,128],	Asch[71] Broadbent[69] Cadogan[72] Ganz[74] Higashi 2004[75] Higashi 2005[136] Kirk[67] McGlynn[77] Osteras[82] Steel 2007[68] Steel 2008[70] Wenger[78]	Feasible: Requires change in routine coding to capture over-the-counter drug use.
Paracetamol 2 [ACR (hip, knee), EULAR (hand, hip, knee), NICE, OARSI]	"IF oral pharmacologic therapy for osteoarthritis 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][127,,128]	ACOVE-1[119,134,]and as adapted (ELSA[119,,134], ACOVE/NH[124], HPCQI[133])	Broadbent[69] Cadogan[72] Higashi 2004[75] Steel 2007[68] Wenger[78]	Feasible: Requires change in routine coding to capture over-the-counter drug use.
Paracetamol 3 [ACR (hip, knee), EULAR (hand, hip, knee), NICE, OARSI]	"Percentage of patient visits for patients aged 21 years and older with a diagnosis of OA with an assessment for use of anti-inflammatory or analgesic OTC medications"[PCPI][68,,69,,72,75,,78,	No other variations	This indicator had not been implemented in any identified studies	*Not suitable: not considered to have been through the same degree of development and testing as most other indicators.*
Paracetamol 4 [ACR (hip, knee), EULAR (hand, hip, knee), NICE, OARSI]	"Percentage of patient visits for patients aged 21 years and older with a diagnosis of OA during which an anti-inflammatory agent or analgesic was considered"[PCPI] [82]	No other variations	This indicator had not been implemented in any identified studies	*Not suitable: not considered to have been through the same degree of development and testing as most other indicators.*
Paracetamol 5 [ACR (hip, knee), EULAR (hand, hip, knee), NICE, OARSI]	"IF a vulnerable elder is prescribed chronic high-dose acetaminophen (≥3 g/d) or a VE with liver disease is prescribed chronic acetaminophen, THEN he or she should be advised of the risk of liver toxicity…"[ACOVE-3][129]	No other variations	This indicator had not been implemented in any identified studies	*Not suitable: considered less feasible to measure from routinely coded data sources – could be implemented with substantial change to coding behaviour.* (con't)

Overarching theme [guidance source]	Exemplar indicator	Reproducibility (studies also developing similar indicators)	References for identified implementation studies	Feasibility assessment
Oral NSAIDs 1 [all]	"If NSAIDs are considered, ibuprofen should be considered for first-line treatment unless contraindicated or intolerant.[†]" [QIGP][122,126,132]	Modifications exist in two implementation studies (Steel[131], Broadbent[68]) to include use of COX-2 selective drugs	Broadbent[69] Kirk[67] Steel 2007[68]	Feasible: Requires change in routine coding to capture over-the-counter drug use.
Oral NSAIDs 2 [all]	"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][67,68,,69,]	Two indicators from ACOVE-3 state that risks from NSAIDs and aspirin should be "discussed and documented,"[129] EUMUSC.net[122,126,132]	Broadbent[69] Steel 2007[68]	Feasible: Requires change in routine coding to capture over-the-counter drug use.
Oral NSAIDs 3 [all]	"Percentage of patient visits for patients aged 21 years and older with a diagnosis of OA with an assessment for use of anti-inflammatory or analgesic OTC medications" [PCPI][68,,69]	No other variations	This indicator had not been implemented in any identified studies	*Not suitable: not considered to have been through the same degree of development and testing as most other indicators*
Oral NSAIDs 4 [all]	"Percentage of patient visits for patients aged 21 years and older with a diagnosis of OA during which an anti-inflammatory agent or analgesic was considered" [PCPI][82]	No other variations	This indicator had not been implemented in any identified studies	*Not suitable: not considered to have been through the same degree of development and testing as most other indicators*

[†] Different sources of guidance offer varying recommendations about the use of specific NSAIDs. In the UK, NICE recommend a standard NSAID or COX-2 inhibitor (other than etoricoxib 60mg) to be co-prescribed with a PPI

Overarching theme [guidance source]	Exemplar indicator	Reproducibility (studies also developing similar indicators)	References for identified implementation studies	Feasibility assessment
Oral NSAIDs 5 [all]	"IF a VE is prescribed an NSAID (nonselective or selective), THEN GI bleeding risks should be discussed and documented…" [ACOVE-3][129]	ACOVE-1 [122,126,132]], HPCQI [119,134,]	Broadbent[69] Ganz[74] Higashi 2005[136] Steel 2007[68]	*Not suitable: Indicator was considered less feasible to measure from routinely coded data sources. Could only be implemented with substantial revision to coding practices.*
Gastroprotection [EULAR (hand, hip, knee), NICE, OARSI]	"IF a vulnerable elder with a risk factor for GI bleeding (aged ≥75, peptic ulcer disease, history of GI bleeding, warfarin use, chronic glucocorticoid use) is treated with a nonselective NSAID, THEN he or she should be treated concomitantly with misoprostol or a PPI." [ACOVE-3][68,69,74,75,78,82,136]	ACOVE-1[122,126,132], ACOVE-3[119,134,] (NSAIDs, and aspirin), QIGP[122,126,132], PCPI[131]	Chodosh[73] Ganz[74] Higashi 2004[75] Higashi 2005[136] Kirk[67] Maclean[76] Wenger[78] Zingmond[79]	Feasible: Should be captured from existing electronic prescribing records.
Specialist Assessment: X-ray [None]	"IF a patient has hip or knee osteoarthritis AND has worsening complaints accompanied by a progressive decrease in activities AND no previous radiograph during the preceding 3 months, THEN a knee or hip radiograph should be performed within 3 months…"[Arthritis Foundation][67,73-75,76,,78,79,,136]	No other variations	This indicator had not been implemented in any identified studies	*Not suitable: Indicator was not considered appropriate for implementation since denominator hard to define; NICE do not recommend routine use of X-rays for diagnosis or referral decisions.*
Specialist Assessment: Referral [EULAR (hand, hip, knee), NICE, OARSI]	"IF a VE has severe symptomatic OA of the knee or hip despite nonsurgical therapy, THEN a referral to an orthopædic surgeon should be made, BECAUSE joint surgery may reduce pain and improve functional status and quality of life."[ACOVE-3][127,,128]	RAND QA [122,126,132], ACOVE-1 [135], and as adapted (ELSA [119,134,]), Arthritis Foundation[124], QIGP [127,,128], EUMUSC.net [131]	Broadbent[69] Ganz[74] Higashi 2005[136] Kirk[67] Osteras[82] Steel 2007[68] Steel 2008[70]	Feasible: It would be feasible to capture the presence of non-surgical therapy indicators in the record, routine data sources cannot be used reliably to assess the need for a surgical opinion.

(adapted from Edwards et al. 2013[107])

Table 2-3: Quality indicators for OA - proposals for primary care implementation

Theme	Proposal for primary care implementation
Holistic Assessment: Pain	Numerator: patients with evidence of pain assessment within the previous 12 months Denominator: all patients with clinical OA
Holistic Assessment: Function	Numerator: patients with evidence of function assessment within the previous 12 months Denominator: all patients with clinical OA
Education	Numerator: patients with evidence of education or advice since diagnosis Denominator: all patients with clinical OA
Exercise 1	Numerator: patients with evidence of exercise advice or physiotherapy referral since diagnosis Denominator: all patients with clinical OA in the hip or knee
Exercise 2	Numerator: patients with evidence of an activity review within the previous 12 months Denominator: all patients with clinical OA
Weight loss 1	Numerator: patients who have a record of weight loss advice within the previous 12 months Denominator: all registered patients with a BMI ≥30 kgm^{-2}
Weight loss 2	Numerator: patients who have a record of weight loss advice within the previous 12 months Denominator: patients with clinical OA with a BMI ≥25 kgm^{-2}
Aids and devices 1	Numerator: patients who are recorded as receiving a referral or assessment for ambulatory assistive devices within the previous 12 months Denominator: patients with clinical OA with evidence of functional impairment
Aids and devices 2	Numerator: patients who are recorded as receiving a referral or assessment for assistive devices within the previous 12 months Denominator: patients with clinical OA with evidence of functional impairment
Paracetamol 1	Numerator: patients with evidence of paracetamol as the first oral analgesic prescribed or advised since diagnosis Denominator: patients with clinical OA with evidence of pain
Paracetamol 2	Numerator: patients with evidence that a suitable maximal dose of paracetamol was tried first Denominator: patients with clinical OA prescribed oral analgesics or NSAIDs
Oral NSAIDs 1	Numerator: patients with evidence of a standard NSAID or COX-2 inhibitor as the first oral NSAID prescribed or advised since diagnosis Denominator: patients with clinical OA prescribed an oral NSAID
Oral NSAIDs 2	Numerator: patients taking an oral NSAID with a documented risk assessment prior to first prescription Denominator: patients with clinical OA prescribed an oral NSAID
Gastroprotection	Numerator: patients prescribed a PPI or alternative gastro-protective agent Denominator: patients with clinical OA prescribed an oral NSAID
Specialist assessment	Numerator: patients who have a record of achievement of all other applicable indicators[‡] Denominator: patients with clinical OA referred for a specialist opinion

[‡]i.e. the other 14 indicators above, depending on applicability of weight and therapy indicators to individual patients

Fifteen indicators were therefore considered to be consistent with the NICE OA guidelines and feasible for implementation in primary care in England. These are shown in the format of definitions of numerator (patients fulfilling the quality criterion) and denominator (patients eligible for the criterion) in Table 2-3. All numerator and denominator statements imply application to patients with a working diagnosis of OA.

The remaining indicators were considered to be less valid (on the basis of the evidence used or consensus approach), inconsistent with NICE guidance (such as the use of X-rays), or less feasible for implementation. The reasons are listed in Table 2-2.

2.4 Discussion

2.4.1 Summary and comparison with previous literature

This review provides an overview of the published quality indicators for the management of OA relevant to primary care in England. It adds to previous work by Hochberg[105] and Strömbeck et al.[106] through identification and appraisal of the indicators within related groups, including use of implementation evidence. A novel aspect to this review was the determination of feasibility for primary care application of the indicators.

The indicators selected in this review were broadly applicable across the different course of guidance on OA management guidance referred to in section 1.4.1. Clinical practice guidelines do not always reflect the state of the art, and quality indicator sets are likewise at risk of becoming superseded by developments in evidence (for example, in relation to concerns about paracetamol use being less safe than previously thought,[139] or the heightened concerns about COX-2 and other oral NSAID use[140]) after the development of indicators. None of the indicator sets were found to have a mechanism to ensure revision in light of emerging evidence. The use of Quality Standards by NICE,[141] which have included OA since June 2015, goes some way towards addressing this issue through use of an annual review of those quality standards to determine whether or not an update

is required. The range of standards for OA mandated by NICE was less broad than the indicators identified by this review.

There were various aspects of guidance which were not clearly reflected by existing quality indicators. Most notably, topical agents have not been the subject of previous indicator development, despite both topical NSAIDs and capsaicin being recommended for hand and knee OA in the NICE guidance.[106] There were many other areas which also have either no indicator or which were not distinctively identified within the existing indicators. These included (i) holistic review – all aspects except pain and function assessment, notably a periodic review, a jointly agreed management plan, and the effect of comorbidities; (ii) education and self-management: the development of a self-management plan, and thermotherapy; (iii) non-pharmacological management: manipulation and stretching, electrotherapy, bracing, joint supports, footwear and insoles; (iv) pharmacological management: topical NSAIDs and capsaicin, and intra-articular injections. In principle, these subject areas were considered suitable for indicator development. As well as a lack of some relevant indicators regarding recommended processes of care, no evidence of any indicators regarding aspects of care that were recommended not to take place ('do not do' indicators) was found. Such indicators might be considered for such aspects of NICE guidance as prescription of topical rubefacients or nutraceuticals, intra-articular hyaluronic acid injections, and referral for electro-acupuncture or arthroscopic lavage.

No evidence of external validity was found for any indicator, so it is unclear whether the use of quality indicators can directly result in quality improvement. Evidence of responsiveness (sensitivity to change) was also not identified so, although an indicator's ability to discriminate between care of a higher or lower quality is assumed, this has not been demonstrated in the literature. Both aspects merit further development: an increased use of patient-reported process measures such as those used by Østerås et al.[82] would be one means of such development, especially if such indicators could be entered directly into the primary medical record to enable routine use in audit.

Implementation of the recommended indicators would generally require some modification to data capture. Prescriptions were a notable exception, since the bulk of prescriptions tend to be electronically generated and the information systems could identify with relative ease the nature of the prescription. This has been successfully used in the QOF[102] for various prescription monitoring indicators such as the use of antiplatelet agents in vascular disease. Many of the remaining indicators were reliant on clinicians taking time to record the full content of the consultation. With appropriate modification to coding behaviour, the selected indicators were considered to be generalisable to general practice across the UK and to international practice where medical records are computerised.

The proposed indicators were designed to be compatible with an episodic review of care for people with OA (such as would occur for long-term condition reviews facilitated by the QOF), so indicators such as the annual assessments for pain and function, an annual review of physical activity and weight advice (where relevant), and an aids and devices needs assessment would need to be sufficiently succinct that they do not trigger clinical disengagement from the process.

It should be noted that the indicators may be most appropriate at a population level:[142] in a population of patients with OA, there could be patients in whom individual indicators would be less appropriate (for example, people who have already been actively self-managing may not need to be referred to an exercise programme, and some patients may present with such advanced disease that an early referral is warranted without necessarily going through the other processes of care beforehand). Population-level thresholds for achievement have not been established.

2.4.2 Strengths and limitations

This review used a sensitive search strategy to identify the available literature on quality indicators. A system of dual-review of abstracts and full-text articles, with dual extraction of data for the included articles and subsequent narrative synthesis, improved the reliability of the review. Alignment to the NICE guidelines and interpretation of feasibility for a primary care context by two

experienced GPs, has improved the applicability of the final indicator set to general practice in England.

There were some limitations in this review. There may be indicators not captured by the search strategy (such as any prior to 2000, and all non-English language indicators). So-called 'grey' literature was also not eligible for inclusion in this review. Given that a number of the identified indicator sets were based upon a thorough evidence synthesis, it seems unlikely that any major themes will have been omitted, though it was possible that some variability within themes may not have been captured. Additional information was sought from the authors of two included articles but the majority of the assessments were made solely on the basis of published information. Quality assessment of quality indicators is not a highly-developed science and skill. The methodological assessment used in this review was a pragmatic approach based on accepted methods of indicator development and a peer-reviewed clinical performance measures checklist.[117] This could be enhanced by establishment of a consensus on quality indicator assessment methodology similar to the COSMIN checklist,[118] such as the Quality Indicator Critical Appraisal (QICA) checklist proposed in 2014 by Jones et al.[143] The subjective nature of some of the assessments made in the review, particularly in relation to feasibility, was a limitation, though it was not considered achievable to create a robust, transparent and reproducible method for making judgments about feasibility and relevance to primary care in England. This could have been strengthened by use of a group of GPs to undertake a formal consensus exercise such as that used by Marshall et al.[144] Acceptability of the indicators to a wider group of GPs was not assessed; an exercise to determine clinical acceptability similar to that undertaken by Marshall et al. would be a potential method of achieving this. The review identifies that some indicators require additional data collection to make routine assessments possible. It could be argued that the proposed extension to routine coding of data should include additional detail. Determination of data collection requirements could again be strengthened through use of a consensus panel. In recommending the synthesised OA primary care indicators, the precision of some of the original indicators has been somewhat diluted to overcome within-group heterogeneity, to maintain concordance with NICE guidance, and to maximise the

feasibility of the indicator implementation. This approach both strengthens the probability of implementation in UK general practice and to an extent undermines the science behind the originally developed indicators. However, the approach taken was in a similar vein to other adaptations of original indicators that have taken place without recourse to further evidence synthesis and consensus development (for example, the inclusion of compound analgesics by Steel et al.,[144] or the interpretation of success criteria by Li et al.[68]).

An additional limitation to the application of the review was the potential difficulty in determining the population to whom the indicators should be applied. Whilst they were applicable to people with a formal OA diagnosis, this would be likely to represent only a part of the population who actually have clinical OA as defined by NICE (section 1.2, Definitions of osteoarthritis). The issue of a practical definition of clinical OA is discussed further in Chapter Three, section 3.5.

2.4.3 Conclusion

In conclusion, a set of 15 indicators considered valid and feasible for implementation in general practice in England have been identified. Structured implementation was considered necessary to achieve uptake of many of these in primary care, given the information capture limitations noted. In Chapter Three, the way in which these indicators might be used within the MOSAICS study is developed and discussed, with their use in assessing quality of care for OA reported in Chapter Five to Chapter Eight.

Chapter Three: Development of an osteoarthritis consultation recording template

3.1 Introduction

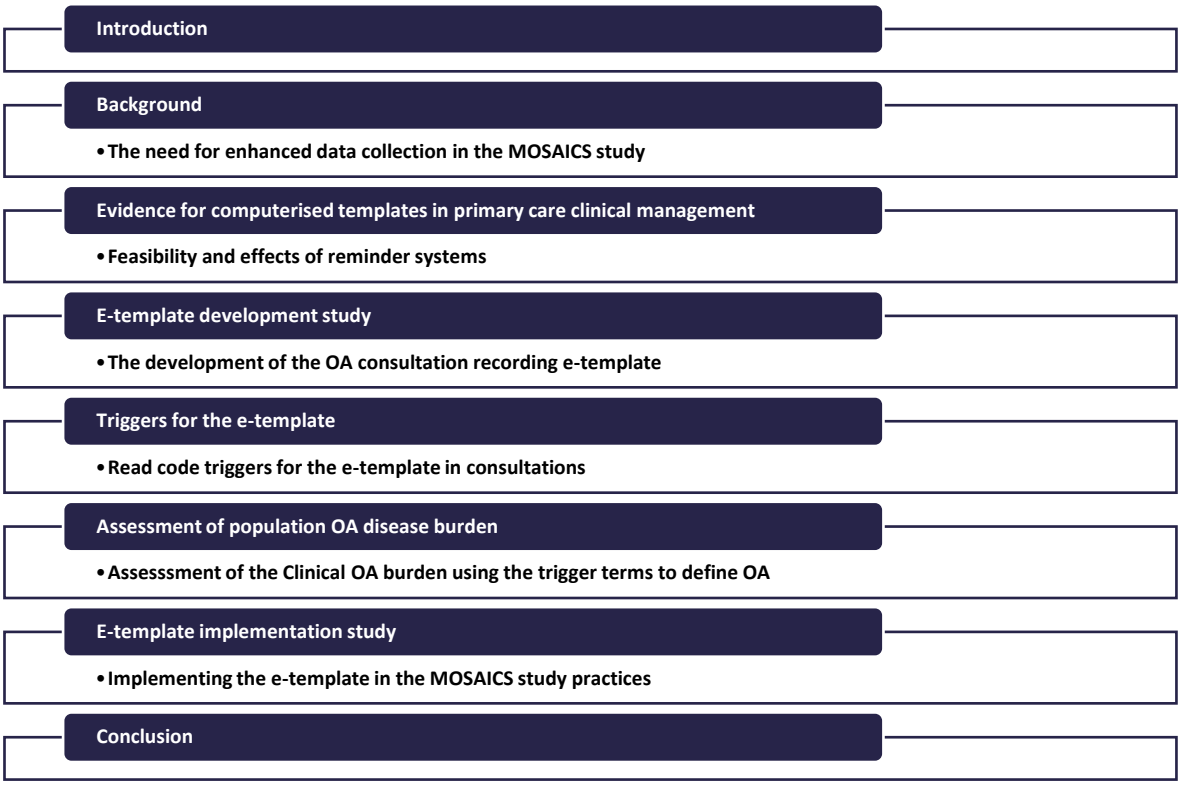
Much of the general practice medical record in the UK is computerised. This has been increasing since the early days of general practice computerisation in the late 1980s, encouraged by policies such as the IM&T DES of the General Medical Services contract 2004.[101] However, the electronic medical record structure retains some similarity with previous paper records, in that only part of the electronic record is coded and therefore easily searchable with automatically retrievable information.[22,,145] Some elements of a consultation would be expected to be recorded in an easily retrievable manner whereas others would tend to require a full analysis of a patient's whole record including the narrative (free-text). Typically, a consultation record would consist of at least one item coded using the Read code system.[146] This could be a symptom, procedure, or process-of-care code, or a disease diagnostic code; the associated narrative rubric is generally of lower utility for routine analysis. Other coded elements may include investigations or referrals. Prescriptions are typically electronically generated and consequently prescribing records can be retrieved relatively straightforwardly. Even if hand-written, a prescription entry should additionally be made in the electronic record.[147] Test results tend to be recorded electronically, though may either result from manual entry of a paper copy of the result or through automated electronic delivery of the result from its originator.

Of the quality indicator outputs from the systematic review[107] (Chapter Two), some could be assessed though use of the medical record as used in standard practice, notably the prescribing indicators. However, standard methods of consultation recording were not considered likely to be adequate to examine other aspects of recorded quality of care for osteoarthritis in the MOSAICS study, as outlined in Table 2-2. Indicators identified as likely to need a change in recording practice were pain and function assessment recording, education provision, exercise advice, aids and devices referrals, and use of (over–the–counter) paracetamol and oral NSAIDs. In order to enhance

the reliability of data collected within the MOSAICS study, it was decided to use an electronic consultation recording template within the study practices. The objective of this chapter is to describe the evidence for such templates in influencing clinical behaviour and show, through a series of developmental studies, the way in which the MOSAICS template was developed, tested, and implemented.

In this chapter, the evidence for the effects of templates on delivery and recording of clinical care, and evidence for the best methods of template development and implementation in primary care are examined. The need for, and development and implementation of, a template for OA in the MOSAICS study is explained. The development and implementation work is presented as a series of developmental studies, outlined in Figure 3-1.

Figure 3-1: Outline of the structure for Chapter Three.



The developmental studies took place over a period of 18 months and, rather than being conducted in a linear sequence, overlapped as determined by the needs of the MOSAICS study timeline. The studies used a mixture of methods and at times a minimal formal evaluation was used. The methods

and results for each component study are discussed individually, with discussion and critical reflection on the potential impact on the content and performance of the template.

3.2 Background: the need for enhanced data collection in the MOSAICS study

The outputs from the systematic review were used to assist the identification of those indicators not identifiable in routinely recorded primary care data and therefore necessary for inclusion in the template.

No published evaluation of the utility of the general practice EHR in the UK at capturing information other than morbidity codes, prescriptions, and referrals was identified. A preliminary scoping exercise was undertaken using the Read code browser[148] and the CiPCA database. CiPCA is a database of primary care consultations from twelve (as at 2007) North Staffordshire general practices trained in, and assessed upon, the quality of recording of consultations and so should represent best practice in recording.[21,22] This exercise demonstrated that some aspects of care were routinely recorded (predominantly, prescribing); referrals were also potentially identifiable from a related database of referrals. However, many aspects of OA care could not be, or were not, recorded. There were no available codes identified for recording assessment of, or advice to use, over-the-counter topical NSAIDs. Only limited codes were identified for pain and function assessment, information provision regarding OA, or paracetamol and oral NSAID use assessment and advice. Even the information that could be recorded through the use of the available codes for pain, function, exercise assessment and advice was not consistently well-recorded; for example, advice regarding weight loss appeared not consistently to have been coded as recorded rates were implausibly low. Physiotherapy referral within consultations was also rarely recorded, implying that although a formal referral may be captured through specific referral data (not interrogated for this aspect of the investigation), other advice to contact physiotherapy (such as private therapy, or physiotherapy triage) was not routinely coded. Some of this information may have been present in the uncoded narrative associated with Read codes but this lacked practicality as an alternative data source due to the lack of sufficiently sophisticated natural language processing software to extract

the relevant information.[149] The alternatives for data collection within MOSAICS were therefore either to undertake an in-depth analysis of each patient's full-text record, or to create a mechanism for improved data recording. Full-text record analysis was not considered to be a feasible method of data extraction for the purposes of comparing intervention and control arms on recorded achievement of quality of care for two reasons; (i) the size of the dataset in MOSAICS (projected eligible study population 30,000 people aged 45 years and over[1]), and (ii) for study practicability, only a small subset of the practice population (those included in the main trial) were to be approached for consent to view their identifiable medical record through a population survey of joint pain with medical record linkage.[1]

To improve the utility of the medical record for the purposes of the MOSAICS study, a computer template was created. This was conceived as a means to improve the routine recording of consultations with patients seeing a clinician due to OA. The template was intended to provide a straightforward means of coding assessments that existing Read codes could not have facilitated (either due to a complete absence of codes or lack of precision) and to prompt the recording of aspects not well-captured even if codes were available (such as for exercise advice).

3.3 Evidence for computerised templates in primary care: feasibility and effects

3.3.1 Background

Although computer templates in primary care have become commonplace, and are routinely used for data capture and structuring of consultations (for example, in the QOF[102]), their use is controversial. For example, Swinglehurst et al.,[150] in a case study of two general practices in England, highlighted a risk of compartmentalisation of the complexities of clinical care into templates, and of framing quality of care through such a potentially reductionist approach. An examination of evidence about e-Health interventions[151] in a wide sense (including data storage, management and transmission; decision support; and remote healthcare) concluded that there was a lack of high-quality evidence to support many of the policy assertions made about some of these

interventions in changing care and expressed specific relevant concerns relating to the time spent documenting patient data as well as the computer becoming a "third person" in a consultation.

The aim of this initial developmental study was to identify what evidence existed about the effects of templates on clinical behaviours, with particular regard to the recording of data, to inform the creation of a recording template for the MOSAICS study. The potential that such a template could act as an additional intervention within the study was recognised at the study inception by the candidate and also by the trial steering committee. In light of this, evidence for templates as a healthcare intervention was also sought, with a view to assessment of the likelihood and degree of any effect on planned trial outcome measures.

3.3.2 Method

A MEDLINE scoping search (without restriction by clinical condition or geographical location) was conducted using NHS Evidence.[§] No systematic reviews specific to the effect of primary care computer templates on consultation recording or clinical management were identified. The search strategy was therefore broadened to consider reminder systems in general, on the basis that templates essentially act as reminders to clinicians to record information or conduct a particular process of care. This identified a relatively recent relevant Cochrane systematic review (Shojania et al. 2009[152]) on the effects of computer reminders on clinical care. This review was used as the basis for identification of the evidence for the feasibility of use and effects on care (processes or outcomes) of the template with any additional evidence published since the review identified through implementation of part of the same search strategy as used by Shojania in MEDLINE. The objective was not to undertake a fresh systematic review or meta-analysis but rather to use a systematic search strategy to identify additional new and potentially important information.

The MEDLINE search strategy used by Shojania et al. and set out comprehensively in their review[152] (Appendix 1, p.64) was used in the NHS Evidence MEDLINE search portal with a publication date

[§]NHS Evidence is a portal through which various bibliographic databases may be searched. https://www.evidence.nhs.uk/ [Accessed 24/09/2016]

restriction from July 2008 (when the previous review search was undertaken) onwards. The resulting citations were downloaded to a citation manager tool (EndNote® X7). The same inclusion criteria as set out in the review were used. The candidate alone determined the relevance of the studies through an assessment, hierarchically, of the title, abstract, and full-text of the identified articles. The search was last updated on 02/09/2015.

Only some results were considered to be of interest. These related to the effects of templates in general, and on the following outcomes:

- Data capture
- Process adherence
- Prescribing
- Test ordering
- Referral to other healthcare providers including therapy professions and secondary care medical professionals

The results are presented in Table 3-1 with the selected relevant conclusions from the Shojania review and additional studies identified subsequent to the evidence search date in that review.

3.3.3 Results

Table 3-1: Evidence for the feasibility and outcomes of reminder systems in clinical practice

STUDY	TYPE	INTERVENTION	DOMAIN	SUMMARY RESULTS		NOTES
Shojania et al.,[152] 2009	Systematic review	Reminder systems – 28 included articles with 32 comparisons. Nineteen based in the USA, 24 outpatient-based.		**Median absolute improvement (IQR)**		
				(Using median outcome from each study)	**(Using best outcome from each study)**	
			All process domains	4.2% (0.8% to 18.8%)	5.6% (2.0% to 19.2%)	Due to multiple studies not reporting a primary outcome, the results were presented as the median outcome with interquartile range and also using the best outcome with IQR to give a range of potential effects on the process-of-care domains. It was unclear if one USA hospital, from where a number of studies had been undertaken, had greater-than-average success rates due to reminder system features or other institutional/cultural factors. Overall, small-to-moderate effects of reminders were seen in the context of considerable heterogeneity between studies (including in the method of reporting) and reminder systems.
			Clinical process (prescription of medication)	3.3% (0.5% to 10.6%)	6.2% (3.0% to 28.0%)	
			Test ordering	3.8% (0.4% to 16.3%)	9.6% (0.6% to 24.0%)	
			Process measures (documentation elements)	0.0% (-1.0% to 1.3%)	2.0% (2.0% to 4.0%)	
			Process measures (other)	1.0% (0.8% to 8.5%)	4.0% (0.8%to 8.5%)	
			Clinical outcomes (dichotomous)	2.5% (1.3% to 4.2%)	n/a	

(con't)

STUDY	TYPE	INTERVENTION	DOMAIN	SUMMARY RESULTS	NOTES
Lecumberri et al.,[153] 2008	Before-and-after study	Automated electronic alerts based on a hospital clinical record database to reduce venous thromboembolism (VTE) risk	Clinical outcome (dichotomous)	No significant effect was seen in year 1 nor for surgical patients or the population overall. Odds for VTE in medical patients was lower by year 2 (OR 0.36, 95% CI 0.12,0.98). Secondary outcomes were the total number of electronic alerts sent (significant in both medical [1.9% increase] and surgical [2.5% decrease] subgroups but not the population overall) and the percentage of alerted patients who received appropriate thromboprophylaxis (significant 15.2% improvement in medical patients but not significant in surgical patients (1.2% decrease) or the population overall (0.3% increase))	
Tamblyn et al.,[154] 2008	Cluster RCT	Automated vs. on-demand customizable computerised decision support for prescribing.	Process measure (prescribing)	No significant effect on the prevalence of prescribing problems (primary outcome) except for reduced therapeutic duplication errors (secondary outcome, OR 0.55, 95% CI 0.33,0.90).	
			Process measure (prompts)	Increase in prompts viewed (secondary outcome: 10.3% of total problems viewed in the automated group vs. 0.9%)	
Lo et al.,[155] 2009	Cluster RCT	Non-interruptive alerts for test ordering relevant to prescribing behaviours	Process measure (test ordering)	No significant difference in recommended test ordering rates (39% of patients in the control group had baseline tests requested, 41% in the intervention group)	

(con't)

STUDY	TYPE	INTERVENTION	DOMAIN	SUMMARY RESULTS	NOTES
Sundaram et al.,[156] 2009	RCT	Clinician-level randomisation to education plus one of two electronic reminder types, with controls receiving education only.	Process measure (screening)	No significant difference in the primary outcome of HIV screening in primary care: control group testing increased from 1% to 1.4% and the intervention group 1.8% to 1.9%. Relevant secondary outcomes included guideline-concordance of testing and reminder adherence. Reminder adherence was stated to be significantly better in the intervention group (11% vs. 5%, $p<0.01$) but it was not apparent that this adjusted for baseline achievement	
Walker et al.,[157] 2010	Cluster RCT	Control practices received education only; intervention received this plus a simple pop-up reminder in all 16 to 24 years old women in Australia, embedded in the EHR.	Process measure (screening)	Improvements in chlamydia testing were seen in the intervention group more than the control (OR 1.3, 95% CI 1.1,1.4)	
Holt et al.,[158] 2012	Systematic review	A review of the effects of patient-specific reminders available in consultations. 42 papers (44 comparisons) included with high heterogeneity. Computer-generated paper-based reminders were included.	Composite of process-of-care measures and clinical outcomes	Summary odds ratio (OR) of 1.79 (95% C 1.56,2.05) favouring reminders was derived from a composite meta-analysis of 44 comparisons (process-of-care measures and clinical outcomes, including screening and vaccination rates, diagnostic tests, blood pressure control, rate of VTE, measures of prescribing quality).	There was high heterogeneity (χ^2=1530.40, I^2=97%) and specific features of computer generated reminder tools that determine the effectiveness were regarded as unclear.

(con't)

STUDY	TYPE	INTERVENTION	DOMAIN	SUMMARY RESULTS	NOTES
Robbins et al.,[159] 2012	RCT	Clinical decision support system for HIV management in the USA. Healthcare providers received interactive alerts (which facilitated care) for half their patients and static alerts for the other half.	Clinical intermediate outcome (continuous measure) Process measure	The primary outcome of CD4 count increases was statistically significantly better in the intervention arm (mean difference 2.0 cells/mm^3/month, 95% CI 0.1,4.0). Secondary outcomes of rate of 6-month suboptimal follow up 20.6 versus 30.1 events per 100 patient-years (p=0.022) and, after a suboptimal follow-up or toxicity alert, time to next appointment (1.71 versus 3.48 months; p<0.001) were better in the intervention group.	
Were et al.,[160] 2013	RCT	Reminders for overdue clinical tasks were integrated into a hospital electronic health record in Kenya. Applied to paediatric HIV care over 6 clinic visits.	Composite process measures	Primary outcome of interest was the number of visits (inclusive) before a recommended action had been fulfilled and documented. In the intervention group, mean time to completion was 77 days (SE 2.4 days), compared to 104 days (SE 1.2 days) for the control group (P<0.001 There was a fourfold increase in the completion of overdue clinical tasks in intervention arm (68% intervention vs 18% control, P<0.001).	Results varied by type of care recommended by the reminder, with reported statistical significance for test requests (HIV ELISA test, chest X-ray, and some other tests including blood chemistry and blood counts) and for referral due to malnutrition Higher rates for commencing antiretroviral therapy in the intervention group did not achieve statistical significance.

(con't)

STUDY	TYPE	INTERVENTION	DOMAIN	SUMMARY RESULTS	NOTES
Kortteisto et al.,[161] 2014	Patient-level RCT	Primary care RCT of 154 reminder rules in the electronic health record in Finland. Intervention patient records received a system for patient-specific reminders to be displayed on opening and closing the patient record; control records were treated according to normal practice – reminders were not displayed but stored in a log file to facilitate comparison with intervention records.	Composite process measures (screening, test ordering, prescribing)	The primary outcome (reminder incidence rate ratio, a composite of change in numbers of reminders in groups over 12 months) showed no significant difference between groups. [incidence rate ratio 1.002 (95% CI 0.995 – 1.009)] In a subgroup analysis, for patients followed up for only 6 months, the number of reminders increased less in the intervention group once confounding factors were adjusted for [incidence rate ratio 0.989 (95% CI 0.978 – 0.9997)].	

3.3.4 Discussion

Since publication of the Shojania review,[152] Cheung et al.[162] (2012) published an overview of systematic reviews of the effects of computer reminders. None of the reviews included in the Cheung overview had a literature search date later than the Shojania review, however. The authors concluded that reminders can have modest beneficial effects on clinical behaviour but that the quality of the review literature on the subject was generally poor.[162]

The studies identified in Table 3-1 suggested that the effects of reminder systems on various process and outcome measures were diverse and not consistently statistically significant. Due to the identified statistical (and methodological) heterogeneity of studies, and the use of composite and secondary outcome measures, it was difficult to draw robust conclusions about the feasibility and impact of reminder systems on clinical practice.

Although there is a growing body of literature regarding template and reminder systems, such systems have been very heterogeneous in their approaches. Some reported systems were more akin to computerised decision support systems, which have not been discussed here. Some were less complex but nevertheless more than simple reminders. A few were just isolated pop-ups, prompting the clinician to act on one specific point.

No evidence was found to indicate how long the benefits of reminders identified in the literature persisted for, though follow-up in some of the studies included in the Cochrane review was 18 months or more. It was, however, notable that in the study by Kortteisto et al.,[161] the primary outcome at 12 months was not statistically significant whereas the secondary outcome at six months follow-up was significant, suggesting that there may be a waning of the initial impact of reminders in practice.

It was not possible to draw firm conclusions from the literature about which aspects of templates and reminder systems have the greatest effect, except that the Cochrane review and other studies agree that computerised systems are better than paper-based, and that systems requiring an active response seem more effective than those which do not, though the latter was confounded by

healthcare provider type. Some aspects of recording, process measures, and prescribing, were seen to improve by modest amounts.

3.3.5 Implications

Given the heterogeneity of the reminder interventions, and variable but generally small to modest effects on clinical behaviour, there was no clear model identified to be followed in the development of the MOSAICS template. The two features identified in the Shojania review as having greater effect sizes (computerisation and active responses) were noted and included in the development model for the MOSAICS template.

It was not clear from the literature how well the template might be expected to function in encouraging data capture, as the Cochrane review showed no statistically significant effect on recording (when using the median of study outcomes), nor was there any definitive evidence about the duration of benefit of the template or any effect of time in reducing template or reminder use.

Although any effects of reminder systems on clinical practice seemed to be small-to-moderate only, it was clear that the MOSAICS study would have to assess the effect of the template itself on quality of care and that it could not be regarded as an entirely neutral intervention. This is addressed in Chapter Seven. It was not considered feasible to avoid a consultation recording template altogether due to the need to capture data from the medical records for all consulters to reduce the risk of recruitment bias in the cluster trial.[163]

3.4 Template development study

3.4.1 Introduction

The RIPCHS has a considerable collective experience in the design and use of templates, arising from the clinical experience of the primary care clinicians who work in the RIPCHS (including the candidate), and from experience in template use in clinical trials, predominantly for trial recruitment purposes.

Given the absence of a template model that could be derived from empirical evidence, the RIPCHS experience was used as a basis for the template development methodology.

For MOSAICS, a recording template was planned to record aspects of care not well-captured by the routine medical record and thus to facilitate recording patterns in three periods (i) prior to the study, (ii) during the template-only phase of the study (prior to randomisation), and (iii) during the MOAC cluster trial phase.

The potential clinical opportunity cost associated with coding items in the EHR was weighted heavily in considerations about the template length. It has been estimated[149] that it can take 30-40 seconds per item coded in the EHR and, although this might be an over-estimate for a structured template such as the one proposed, it would nonetheless represent an unacceptable intrusion into a standard general practice consultation were this to be accurate.

The aim of this developmental study was to design a consultation recording template for the MOSAICS study that would facilitate capture of information within consultations for OA whilst retaining acceptable simplicity and brevity so that clinicians would continue to use it for at least the duration of the study.

3.4.2 Method

Initially, it was necessary to identify which quality indicators could be assessed through data captured routinely by the general practice EHR and which would need to be included in the consultation recording template. No previous defined set of Read codes to determine achievement of OA assessment quality indicators was identified. Therefore, the candidate used the Read browser[148] to determine the codes that mapped to the quality indicators (see Appendix E.1). A search through relevant sections (assessment, processes of care) of the Read hierarchy and of key text words, supplemented by clinical knowledge, was used to identify codes associated with the process-of-care concepts. All codes were searched for through use of keywords or word stems (such as, for the pain and function domains, “pain”, “function”, “ability”, “gait”, “mobility”, “aid”) and inspection of the Read hierarchy for relevant terms. These were investigated in the CiPCA

database[22] for the year 2008 to determine which codes had actually been used in general practice. These were compared with evidence from prior assessments of the quality of OA care in general practice[68-70] to judge the apparent completeness of recording of quality measures through Read codes. This judgment was used to determine which aspects of the quality indicators could be measured by use of Read-coded data without further intervention. From this, a list of aspects of care that were not routinely captured was created. A sample template was developed which was taken through an iterative process of development within the MOSAICS team and other experts-by-experience within the RIPCHS until the MOSAICS team (predominantly, the candidate and his supervisory team, the Chief Investigator, the MOAC intervention lead GP, and the Comprehensive Research Network (CRN) health informatics staff) considered the template to represent the best balance between information inclusiveness, technical feasibility, and clinical acceptability. The way in which the EMIS clinical system presented a template, with a 10-item limit per page, was relevant to the acceptability discussions: the experience of others has been that clinicians dislike using guidelines that do not fit on a single screen.[164]

Once the domains to be included in the template had been agreed, a set of potential response options was agreed following an iterative process of development, discussion, and refinement through the MOSAICS team, including assessment of potential responses through pre-existing Read codes. To maintain homogeneity of responses within the trial, it was agreed within the MOSAICS team that the template responses should be determined from a short pick-list of options, except where a continuous response was required (weight measurement). The homogeneity was considered necessary to allow rule-based dichotomisation into achievement or non-achievement of the quality indicators. This approach maintained the need for the clinician to provide an active response to the template, consistent with the conclusions of the Cochrane Review[152] that point-of-care reminder templates had the greatest effect where an active response was required.

The resulting template was tested using an alternative clinical system (INPS *Vision*) in the candidate's own practice and in the (EMIS-using) practice of the lead GP for the MOAC intervention.

Although the system in the candidate's practice had some interface differences, key elements (triggering and feasibility of completion) could still be assessed.

Once agreed, the CRN health informatics staff created the template within the EMIS clinical system, which was the system used by all trial practices (as a prerequisite for MOSAICS study participation). This used the flexibility of the EMIS system to create bespoke Read codes for use within the study practices, meaning that all template responses could be tagged with a retrievable code.

3.4.3 Results

Prior routine code use

The total registered population of CiPCA-contributing practices was 104,965. Table 3-2 shows the number of patients in whom information relevant to the template domains was recorded. Very low frequencies of recording of information in all domains were identified. It is not known what proportion of patients actually have these care processes performed (rather than recorded) in a year and therefore it is not clear how well the records reflect the care delivered. Referral information was considered likely to have been available in MOSAICS (not captured by the CiPCA database). For the purposes of the MOSAICS study, the levels of recording of information about quality of care other than referrals seen here were considered inadequate to allow comparison of quality indicator achievement between trial arms. The template development therefore continued with decisions about which domains to include.

Table 3-2: Frequency of Read codes relevant to selected OA process-of-care measures

Domain	OA patients in whom relevant consultation Read code recorded *n* (%)
Pain assessment	7 (0.2)
Functional assessment	16 (0.3)
Information provision/education	1 (<0.1)
Weight[a]	14 (0.4)
Weight loss advice or referral	1 (<0.1)
Exercise grading	1 (<0.1)
Exercise advice/referral[b]	2 (<0.1)
Physiotherapy referral[b]	3 (0.1)
Occupational therapy referral[b]	0 (0)
Specialist assessment[b]	0 (0)

[a]this was likely substantially to underrepresent the total frequency of weight measurement due to the way CiPCA data is extracted from the clinical system; [b]these domains examined only Read-coded referrals in consultations: a separate CiPCA-like referrals archive was not interrogated and so the referrals estimate here is considered likely to be a substantial under-estimate.

The quality domains considered necessary for inclusion in the MOSAICS study are shown in Table 3-3. These addressed the most essential holistic assessment and core management for OA as recommended in the NICE guidelines (2008) as well as the recommended first-line pharmacological management strategies. The planned data source is shown (following review of the CiPCA analysis shown in Table 3-2).

Table 3-3: Quality indicator domains for template-derived recorded management of OA

Quality indicator	Indicator origin	Assessment data source(s)
Assessment of pain	Systematic review	Template
Assessment of function	Systematic review	Template
Weight/BMI record	[Prerequisite for weight loss advice indicator]	Routinely recorded or template
Assessment or advice about paracetamol use	Systematic review	Template
Prescription for paracetamol	Systematic review	Routinely recorded prescriptions
Assessment or advice about topical NSAID use	NICE guidance	Template
Prescription for topical NSAIDs	NICE guidance	Routinely recorded prescriptions
Evidence of education or advice for OA	Systematic review	Template
Weight loss advice for people with peripheral joint clinical OA and a BMI ≥25kgm^{-2}	Systematic review	Template
Consideration of physiotherapy referral	Systematic review	Template
Exercise or physiotherapy referral	Systematic review	Routinely recorded codes
Assessment or referral for assistive devices (ambulatory or non-ambulatory) – intended for people with functional impairment	Systematic review	Routinely recorded codes or template
Co-prescription of gastroprotection (for people prescribed oral NSAIDs)	Systematic review	Routinely recorded codes
Completion of all template indicators above for people receiving specialist referral	Systematic review	Template

Template domains

An initial draft template (Appendix E.2), based on the quality indicators identified from Chapter Two included all the proposed indicators from the systematic review with the exception of the prescribing indicators, which were considered possible to derive from the prescriptions records. This draft was put forward for discussion amongst the wider trial team and health informatics specialists.

A pragmatic decision to limit the template to a single page was taken, after discussion between the candidate, his primary supervisor, and the CRN informatics staff, due to concern that a template which extended beyond one page would be less clinically acceptable, especially to the GP staff. The template needed to capture the core aspects of management listed in the NICE OA guidelines.[5] These were considered to represent selected aspects of holistic assessment (pain, function), and the core interventions of advice and education, exercise, and weight loss where relevant.

In addition to the indicators from the systematic review, an indicator regarding the use of topical NSAIDs as a relatively safe pharmacological option was included to be consistent with the NICE guidelines for initial analgesic use. The final template recording domains selected were:

1. Pain assessment
2. Function assessment
3. Weight
4. [System automatic calculation of body mass index, using last height record carried forward]
5. Record of prior use of or recommendation for paracetamol use
6. Record of prior use of or recommendation for topical NSAID use
7. Information provision
8. Advice regarding weight [for people who are overweight: BMI ≥25 kgm^{-2}]
9. Advice regarding exercise
10. Advice regarding physiotherapy referral

The order in which the domains appeared was agreed between the candidate and the academic GP leading the model OA consultation (MOAC) intervention for the MOSAICS study.[1] The order above was considered to be the best fit with the MOAC structure.

The restriction to 10 items inevitably meant that some aspects of care, which might have been desirable to collect in order to understand the primary care of OA better, were not included in the final design. Those not considered to be priorities for individual inclusion in the template were:

- Assessment of mobility
- Assessment of current exercise activity
- Risk assessment (oral NSAIDs)
- MED3 incapacity for work or fit note certification
- Over-the-counter drug use
- Specific contra-indications to drug therapies (paracetamol, NSAIDs in particular)
- Referral to dietician**
- Referral to weight management programme**
- Referral to exercise programme** (other than physiotherapy)
- Referral to occupational therapy**
- Referral to podiatry**
- Referral to orthotist**
- Referral to secondary care** (rheumatology, orthopædics, pain clinic)

Template response options

A clinician could complete all, some or none of the domains either during or after a consultation. It was possible to press 'escape' to bypass the template. It was also possible to move between fields without entering data and file an incomplete template response to the medical record, in line with its use as a recording aid rather than a minimum data set. In all cases, a bespoke code was designed to be written to the record indicating that the template had successfully been triggered, whether or not template data had been entered by the clinician.

A quality indicator needed to be achieved only once in a time period rather than each time the template was triggered, along similar lines to the responses to templates used in the QOF.

**Actual referrals (rather than consideration) could be measured through the referrals data collection rather than the template

i. Pain and function

The potential for use of a numerical rating scale (NRS) for pain and function assessment was considered and discussed. However, since the use of such a scale in primary care has known deficiencies,[165,166] and the template could be completed after a consultation rather than strictly contemporaneously, it was considered that the use of a NRS may give a spurious sense of accuracy regarding pain assessment. Whilst this would not have had a significant impact upon the research project (since the medical record was not being used to track patients' pain over time), it may have leant an unwarranted apparent validity to the information in the medical record. This decision is consistent with subsequent evidence that there is only moderate agreement between the EHR information about pain and that actually reported by patients in a survey.[167]

Therefore pain and function was recorded using a simple ordinal scale:

- Pain: None, mild, moderate, or severe
- Function impact: Not limited, mild limitation, moderate limitation, severe limitation

Any of these responses were considered to meet the quality indicators for assessment of pain and function. If the responses were not considered appropriate, or the pain or function status was unknown, the whole domain(s) could be skipped with no data entry, in which case the quality indicator would not have been met on that occasion.

ii. Weight and BMI

A weight record needed to be added to the template for an updated BMI calculation to be made. If no weight record was made, the last weight and BMI record in the system would be displayed in the template. For brevity, it was assumed that a height record would be present already and that it would not have changed.

iii. Paracetamol and topical NSAID use

As this template was intended for use in both prevalent and incident cases of osteoarthritis, the responses to the assessment of paracetamol and topical NSAIDs needed to account for the fact that patients may have previously tried these agents, or that their pain may have become too severe for

it to be considered clinically sensible to revert to plain paracetamol or topical NSAIDs. It also needed to capture patients' independent use of such drugs, for example on an over-the-counter basis. Thus, this domain was distinct from the indicator assessment through analysis of electronic prescription records. For each of these two pharmacological domains, the following possible responses were included:

1. Tried full dose [i.e. at some point the patient has tried a clinically appropriate maximum dose of the agent]
2. Advised full dose [i.e. the clinician recommends this, either to be obtained over the counter or supplied through prescription]
3. Declined full dose [i.e. the patient declines to try the recommended treatment]
4. Not appropriate [i.e. it was not considered clinically appropriate, for example due to the ongoing use of much more potent agents, or comorbidities causing concern about the appropriateness of such drugs]
5. Unknown [i.e. the clinician entering the information does not know about paracetamol use – for example, when the consultation was over and the patient has left]

The first four of these were considered to demonstrate achievement of the quality indicator.

iv. OA information given

This intention of this domain was to capture the use of the OA guidebook, an integral part of the MOAC. However, since the template needed to be identical for the intervention and control arms of the MOSAICS study, a generic OA information domain was created. Since GPs give information to patients in different ways, both verbal and written methods were considered possible, though for the sake of simplicity, it was assumed that written information would only be given in conjunction with some verbal information:

1. Verbal + written
2. Verbal only

3. Not this time [i.e. the information may have been given before, or was planned for a subsequent consultation]
4. Not appropriate [e.g. the patient was unable to comprehend the information]

All except '*not this time*' were considered to demonstrate achievement of the quality indicator.

v. Weight loss advice

This was only relevant for people defined as overweight, i.e. with a BMI ≥25kgm^{-2} – and for such patients, it was considered that a similar approach, and responses, to the provision of OA information should be taken (see above).

vi. Exercise advice

This presumes that some assessment has been made of a patient's prior exercise habits, although the template does not capture this. There was considerable similarity to the other advice domains:

1. Verbal + written
2. Verbal only
3. Not necessary [i.e. the patient was already considered to be exercising adequately]
4. Not this time [i.e. the information may have been given before, or was planned for a subsequent consultation]
5. Not appropriate [e.g. the patient was unable to exercise]

All except '*not this time*' were considered to demonstrate achievement of the quality indicator.

vii. Physiotherapy referral advised

This domain captures discussion about physiotherapy referral, and overcomes the variation in physiotherapy access (such as use of physiotherapy triage telephone advice lines, private physiotherapy, etc.). The possible responses were:

1. Offered referral

2. Not necessary [e.g. for patients already exercising adequately or using other sources of information such as the Arthritis Research UK information leaflets[168] as a basis for structured exercise]
3. Not this time [i.e. the referral may have been made before, or was planned for a subsequent consultation]
4. Not appropriate [e.g. the patient was unable to exercise]

All except '*not this time*' were considered to demonstrate achievement of the quality indicator.

The final agreed template (version 3) is shown with screen captures in Appendix E.3.

Preliminary testing in the candidate's practice

The template was tested over 46 patients in four months. It was considered straightforward to complete, taking approximately an additional two minutes to populate all the domains. This caused difficulties if OA was not the primary reason for consultation and so the frequency of completion of the template was not as great as desired. No formal feedback was received from the other (EMIS) practice in which the template had been installed for preliminary testing purposes but ad hoc feedback suggested that the template was considered feasible. In light of the candidate's experience of the template, in conjunction with ad hoc feedback from the other (EMIS) practice in which the draft template was tested, it was determined that no further changes to the template were required.

3.4.4 Discussion

The evidence for template development applicable to primary care was not sufficiently strong to be able robustly to state clear principles by which a template should be developed and implemented, other than for the use of computer-generated reminders (computerisation being regarded as a 'given' in a paper-light NHS[101]) and for reminders to require an active response from the clinician. This left considerable scope for opinion-based development of the template, based upon the clinical experience of the GPs contributing to the template development and the

considerable experience the RIPCHS held collectively in developing templates for use in research studies.

The main purpose of the template was to assist in the capture of information relevant to the core aspects of management identified in the 2008 NICE guidance: assessment, education, exercise, and weight loss advice where relevant. Inclusion of a template item to record consideration of physiotherapy referral was also identified as necessary to estimate use of 'prescribed' exercise. Recorded actual referral may not adequately capture clinical conversations about physiotherapy, and use of self-referral schemes such as Physio Direct[169] potentially limited the completeness of the primary care record of use of physiotherapy services.

There was inevitably some tension between the desire to include as much information as possible in the template and to maximise the possibility of clinicians completing it in consultations. The exact length of the template was the subject of much debate, and even once the one-screen size had been agreed, further discussion about the relative priority of different domains was required. A range of the aspects of holistic assessment referred to in the 2008 NICE OA guidelines[5] were also excluded for the sake of brevity. Although the core management strategies were captured by the template or routine medical record, there are some of the wider aspects of management not captured, such as use of assistive devices, thermotherapy, shock-absorbing footwear, supports and braces and transcutaneous electrical nerve stimulation (TENS). Identification of onward referral (except for physiotherapy) was not included in the template and was therefore dependent upon the standard recording methodology in the primary care record. This may not identify all referrals made due to different referral methodologies including the use of electronic referral services such as 'Choose and Book,' as well as different recording strategies between practices, and was not able to capture discussion without actual referral.

Assessment of pain and function was infrequently recorded in the routine records. The approach to assessment and recording adopted in the template was selected for ease and brevity, and due to the fact that it was the assessment of the pain and function itself which was required to pass the

quality indicator for each of these domains rather than valid and reliable longitudinal measures of pain and function. A feasible method of recording such an assessment of pain and function in musculoskeletal conditions would be a desirable change to introduce to primary care, though it was not considered to be within the scope of this project. The validity of some of the other template items could also be questioned, especially if clinicians completed the template at the end of a consultation rather than face-to-face with the patient during the consultation, as this approach risked information or recording bias. However, clinicians were not under any obligation to complete all domains of the template, and 'Unknown' was a possible response option (although it would not have achieved the quality indicator).

It was not possible from the template data to identify the quality of verbal or written advice. There was a risk that response could be ticked without any meaningful engagement with patients. This problem was not unique to this template, as the same charge may be levelled at various information and advice or support indicators in the QOF.

It was not planned that all of the template had to be completed at every relevant consultation and, just as with aspects of chronic disease reviews relevant to the QOF[170] such as diabetes or asthma, elements could be completed at different consultation points within a period by any appropriate clinician (usually taken to be a GP or practice nurse).

The template response options were constrained to give a pick list from which the pass and fail criteria for the quality indicators could straightforwardly be defined. Inevitably, some of the nuances from a well-completed full-text response will have been lost through this approach but it was considered that the benefits of easily-retrieved and standardised information outweighed this loss. Although some increased granularity of response options (as noted by Hersh et al.[171]) might be desirable, there was a risk that this would impact on template completion through increased complexity, though no empirical evidence was available to guide this. The choices made were considered to be pragmatic in light of previous RIPCHS experience of template use.

The criteria for passing or failing the indicator might also be considered debatable. The responses chosen were considered to reflect the most common clinical responses that might be given and so were all included as options rather than increase the risk of the template being left blank for some domains.

Chapter Five examines the routinely-recorded care for OA in the MOSAICS practices, which did not have the formal training in recording provided to CiPCA practices, and so might be more representative of typical general practice. Chapter Six examines the use of the template in measuring quality of care in MOSAICS.

3.4.5 Implications

Not all domains that would ideally have been captured by the template were included. However, it was not considered likely that the routine record would have differential recording patterns between intervention and control arms of the MOSAICS study, so resultant information bias was not considered likely.

Quality indicator success was defined by predetermined response options, which may not always have suited a clinician's consultation style – for example, a clinician might have felt that there should have been more flexibility in the responses and so not completed the template for some or all domains, or alternatively if the template were to be completed at the end of a consultation after a patient had left, some relevant information may not have been obtained. It seems likely that the template would generally act as a positive aid to recording, however, especially given the very low levels of coded information about many of the indicator domains identified in the routine record.

3.5 Triggers for the template and a Read code definition of 'clinical OA'

3.5.1 Background

Reminders that present themselves actively to clinicians appear to be more effective than those which are passive and have to be initiated by the clinician.[162] Therefore, in order to maximise the possibility that a clinician would use the template to record a consultation with a patient regarding

OA, a list of candidate codes for OA was derived and subsequently used to trigger the template actively in a consultation. These codes were considered to be those a GP might use for OA, whether or not a formal diagnosis had been made. Symptom-based codes were included as candidates for OA, since it is known that OA is the most common cause of joint pain in the MOSAICS study target age group of 45 years and over.[172,,173]

The Read code (version 2) system is a hierarchical list of 5-byte alphanumeric codes.[146] Disease codes for OA are mainly listed under *Chapter N.... Musculoskeletal and connective tissue diseases*, though not exclusively under *N05.. Osteoarthritis and allied disorders*, with additional codes describing osteoarthritis symptoms being grouped in this musculoskeletal disease chapter. There were also symptom-based codes, which have potential relevance to OA, in *Chapter 1.... History / symptoms* and *Chapter R.... [D] Symptoms, signs and ill-defined conditions*.

The aim of this developmental study was to identify a set of Read codes to define a working diagnosis for OA (hence "clinical OA") to trigger the template in consultations, and to act as the MOSAICS study definition of clinical OA for the purposes of comparing intervention and control arms on recorded quality of care.

3.5.2 Methods

A list of potential codes to trigger the template was created, drawn from the NHS Read code browser by the candidate.[148] These codes were selected on the basis of relevance to the MOSAICS study (excluding codes relating to spinal OA and to shoulder and elbow joint pain). All codes specifically referring to non-spinal OA were included. Other symptom-based codes, whether or not they fall into a Read symptom or examination code (Read chapters 1 and 2) or disease classification (Read chapters A-Z) within the hierarchy, were included if they related to peripheral joint pain that might have implied clinical OA.

Once a shortlist of codes had been created, a panel of six academic GPs in the RIPCHS (including the candidate), all with an interest in musculoskeletal research, were independently asked to decide whether or not the selected codes should be included in the list of candidate codes for OA.

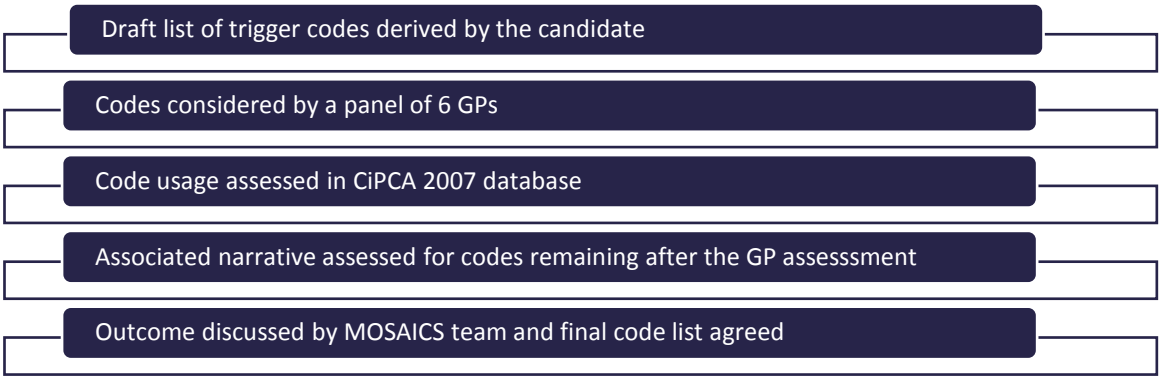
This was done by sending a list of codes in a Microsoft® Excel® spreadsheet and asking the GPs to determine if the codes represented possible clinical OA ("*The task: - to agree a list of Read codes/terms that will be used in the MOSAIC studies to define OA. All entrants will be aged 45 and over, so there is no need to be concerned about adolescent anterior knee pain, for example. Chapters 1, 2, N and R contain codes of possible relevance. Not all child codes are listed (only where considered especially relevant) and some site specific codes will not be appropriate to the study. The full list of child codes can be seen via hyperlink from the child codes column if needed…*"). Each respondent was blind to the others' responses at the time of completion. Codes were to be included if there was a simple majority in favour. In case of a tie between recommendations to include and exclude, the code was retained. For consistency, some of the free text responses were adapted from the original ("Yes?" became 'unsure', "Why not…?" became 'include').

Read coded consultation data were extracted from the CiPCA database[22] for the year 2007 to determine whether or not the Read codes from the GP assessment had actually been used in clinical practice. Codes other than *N05.. Osteoarthritis and allied disorders* and its child codes were discarded from the trigger code list if they had not been recorded in CiPCA to simplify the technical aspects of template installation in the practices.

In order to assess the specificity of the remaining non-N05.. codes to be potentially included in the definition of clinical OA, a random sample of 850 consultations (10% of the clinical OA consultations) held in the CiPCA database (in patients aged 45 years and over) was inspected to determine whether the consultation text provided any evidence that the clinical diagnosis was likely to be OA or not. This was necessarily rather subjective, as it depended on a post-hoc analysis of another clinician's contemporaneous notes. Evidence that the consulting clinician considered the diagnosis to be OA (through use of terms like 'OA', 'crepitus', 'degenerative', 'wear and tear', or reference to X-ray evidence) or likely not to be OA (e.g. evidence of substantial early morning stiffness, other inflammatory symptoms or signs, neurological symptoms) was used to allocate the consultation to OA or not OA. The outcomes were discussed between the candidate, the MOSAICS

team, and the CRN health informatics team. Codes considered to have a low specificity for clinical OA were removed.

Figure 3-2: Summary of template trigger code list determination



3.5.3 Results

Trigger code list selection and GP panel assessment

There were 348 Read codes in total that were assessed for inclusion, 23 higher-level codes and 325 codes that were children of the other codes. These were screened initially by the candidate alone. Formal OA diagnoses (N05.. and child codes) represented 99 of these 348. Of the remainder, 41 codes (the other 22 originally identified higher-level codes plus 19 of the most relevant child codes) were considered by the candidate sufficiently relevant for the other clinicians to consider individually; other child codes were considered in conjunction with their parent code. There were three identified 'local codes' created by practices within CiPCA 2007 that were also included in the template trigger code list. As these were determined by practices, their use is not generalisable, however.

The list of candidate codes derived from the Read code browser is shown in Appendix E.4, page 381. Also shown in the same appendix are the responses from the six RIPCHS GPs regarding proposed inclusion.

Of the 42 codes or clusters of codes (i.e. a code with all child codes included) in the shortlist, the agreement about inclusion or exclusion is shown in Table 3-4 (any positive statement about inclusion as an acceptance of inclusion, and any other statement as a rejection of inclusion):

Table 3-4: Inter-rater agreement on candidate code inclusion

Inclusion		**Exclusion**	
100% agreement to include	3 codes/code clusters	100% agreement to exclude	5 codes/code clusters
83% agreement to include (5 of 6 raters)	3 codes/code clusters	83% agreement to exclude (5 of 6 raters)	9 codes/code clusters
67% agreement to include (4 of 6 raters)	5 codes/code clusters	67% agreement to exclude (4 of 6 raters)	11 codes/code clusters
No overall agreement			
(3 raters for and 3 against) [codes retained]	6 codes/code clusters		

Where there was no overall agreement, codes were retained, which was consistent with the recommendations of the two GPs most closely involved with the MOSAICS study

Testing in CiPCA 2007

The codes resulting from the GP assessment exercise were tested in CiPCA 2007. All N05.. codes were defined as required for inclusion but of the remaining codes, those not recorded in 2007 were removed from the list of trigger terms. Of the codes from the GP consensus (Appendix E.4), those removed at this stage were *1M12. Anterior knee pain* and *2G26. O/E - hands - Heberden's nodes*. The presence or absence of the codes in CiPCA is shown in Appendix E.5, page 383.

Specificity assessment for non-N05.. draft clinical OA codes

A random sample of 850 consultations in the CiPCA database drawn from 8445 (10%) Read-coded consultations using codes identified in Appendix E.5 were used to determine if the consultation was likely to reflect clinical OA or not. The process identified only a small proportion of consultations (104/850, 12.2%) were regarded as representing “true OA” by the candidate; 56/850 (6.6%) were considered not to be OA, and the great majority (81.2%) could not be allocated with any useful degree of certainty. This was found to be due to (i) constraints on the length of the consultation text extracted (EMIS LV allowed only limited narrative text associated with a problem title to be downloaded), (ii) the use in some cases of Read codes that were wholly unrelated to the free text, and (iii) the use of a Read code linked to prescription or administrative data with no additional clinical information. In some cases, it was unclear what the underlying working diagnosis was as the

narrative text appeared complete but was too vague to determine the diagnosis. A systematic assessment of diagnostic information in the MOSAICS practices, considered better to represent usual coding practice, is described in Chapter Four.

The codes that had not been removed in the GP assessment or CiPCA 2007 inclusion assessment exercises were presented to the MOSAICS team by the candidate and a final determination was made on the basis of the study needs and feasibility for linkage of the template to the trigger codes. All *N05.. Osteoarthritis and allied disorders* codes were retained. The list of joint pain codes taken forward as template triggers is shown in Appendix E.6, page 386: those codes retained were considered to represent an acceptable balance of feasibility and inclusiveness and were recorded as having been used in the CiPCA database.

3.5.4 Discussion

The risk of selection bias[174] was reduced through the use of a reasonably wide definition of clinical OA. Clinicians had the option to avoid the template by simply pressing the 'Escape' key but it was harder to avoid the template triggering by avoiding coding OA or joint pain altogether. It was expected that this would reduce any potential for selective inclusion (bias) of OA patients in the population-level analysis, as might have happened if only N05.. codes were used to define OA. The internal validity of the trial was therefore considered to be strengthened by this approach.

However, given the inclusion of terms not specific to OA, one would expect there to be some problems with misclassification, as patients with non-OA joint pain would under this approach be classed as having clinical OA. Depending on the extent of variation in nature of the coding of OA and joint pain within practices and between intervention and control arms, this could mean that the population with actual clinical OA may not be comparable between arms. One potential consequence would be that the proportion recorded as achieving the quality indicators may be different between arms due to the misclassified additional cases. Depending on the exact levels of baseline misclassification and any changes in coding behaviours throughout the trial, this could give rise to misclassification bias.[174] Overall, however, since all patients with peripheral joint pain or OA

were designated as 'clinical OA,' misclassification was considered to be a bigger concern for external validity of the study rather than an issue of bias. As noted in 3.7 below, the feedback from the template implementation showed variable use of, and attitudes toward, the template and so the risk and possible effect of bias from misclassification was hard to assess. In terms of the external study validity (generalisability of the findings), some caution would have to be exercised in applying the findings to populations of people with disease thus defined, as it may not all be OA.

Coding of OA and joint pain in the baseline phase of MOSAICS, before the template installation, is described in Chapter Four, with an assessment of the degree of diagnostic misclassification.

3.6 Assessment of population OA disease burden

3.6.1 Background

The aim of this developmental study was to identify the effect of inclusion of peripheral joint pain codes in people aged 45 years and over on the consultation prevalence of OA. This was also to be used as a guide for recruitment planning for the MOSAICS study.

3.6.2 Methods

The codes identified in section 3.5.3 were used to define clinical OA. The data recorded in CiPCA were used to identify the proportion of the registered population consulting in a 12-month period (2007) with one of these codes recorded in their medical record. The registered population aged 45 years and over was used as the denominator population, in line with the NICE definition of osteoarthritis.[5]

Patients recorded as consulting in the whole of the CiPCA database (12 practices) in 2007 were stratified by age and sex. The results from CiPCA were used to create a "typical" representative practice of 10,000 patients. Also calculated was the expected annual consultation prevalence for OA in this representative practice, according to the wider definition (an N05.. OA diagnosis and the selected joint pain codes considered to represent clinical OA, from section 3.5.3) and the more narrow definition (Read code N05.. *Osteoarthritis and allied disorders* only).

3.6.3 Results

The total population of people aged 45 years and over in CiPCA 2007 was 46,927. The annual crude consultation prevalence rate, with clinical OA defined as either a formal OA diagnosis (N05..) or the selected joint pain codes as described, was estimated at 9.2% in patients aged 45 years and over (4.1% of the total all-age population). For a formal diagnosis of OA using N05.. *Osteoarthritis and allied disorders* Read codes only, the consultation prevalence was estimated at 4.1% in patients aged 45 years and over, and 1.8% in the population as a whole.

Thus, in a representative practice of 10,000 patients typical of the population contributing to CiPCA, one might expect to see the number of cases displayed in Table 3-5.

Table 3-5: Expected numbers of patients consulting annually with OA, per 10,000 population

Registered population	**Males**	**Females**	**Total**			
Age <45 years	2788	2713	5501			
Age ≥45 years	2116	2312	4499			
All ages	4905	5095	10000			
Consulters with OA		**Clinical OA**			**Formally diagnosed OA**	
	Males	**Females**	**Total**	**Males**	**Females**	**Total**
45-54	35	43	78	13	21	34
55-64	47	66	113	18	32	50
65-74	43	64	107	17	31	48
75-84	27	56	82	10	27	37
85+	7	20	27	3	10	13
Total	159	249	408	61	121	182
% total population	3.2	4.9	4.1	1.2	2.4	1.8
% population 45 years and over	7.5	10.8	9.2	2.9	5.2	4.1

3.6.4 Discussion

These prevalence estimates were similar to other estimates from research studies. For example, Jordan et al.[21] determined OA consultation prevalence rates in people 15 years and over for 2001, using the N05.. *Osteoarthritis and allied disorders* OA definition, of 164 to 276 per 10,000 (using data from the GPRD, the RCGP Weekly Returns Service (RCGP-WRS), and CiPCA).

Subsequent to this analysis, another study has identified similar annual prevalence figures to those found here in both the CiPCA database for the year 2010 and in a similar consultation database in Lund, Sweden.[23] This analysis used the both a narrow (N05.. and the International Classification of Disease v.10 [ICD-10] equivalent) and the clinical OA definition, in people aged 45 years and over.

Using data from primary care, and the using the clinical OA approach (diagnosed OA plus relevant symptom codes), consultation prevalence rates of 1074 per 10,000 in CiPCA and 967 per 10,000 in the Lund database were identified. These estimates were reasonably similar to the 9.2% consultation prevalence suggested from the CiPCA 2007 data.

The extensive use of symptom-based codes, especially given the relatively low prevalence of osteoarthritis as a formal diagnosis (N05..) compared to the prevalence of OA using the clinical OA definition, might imply a lack of confidence or reluctance to diagnose osteoarthritis. This contrasts with GP behaviour in other clinical areas, such as asthma and hypertension, where there may have been some cases of over-diagnosis of these conditions (albeit with other cases remaining unidentified),[175-177] tightened up by greater stringency in diagnostic requirements in the QOF (such as by a requirement to diagnose asthma with evidence of abnormal serial peak flow readings), or in national guidance on hypertension emphasising home-based or ambulatory blood pressure monitoring.[178] Were a core set of codes introduced for OA into primary care (as a contractual requirement), one might expect to see a rise in the use of these core codes, given the incentives for higher prevalence in the QOF. This may be seen in the same light as other chronic disease registers where case finding and list-cleaning efforts have taken place. It should be noted, however, that not all conditions included in the QOF have seen an increase in prevalence as a consequence. One study of temporal trends in the prevalence of depression[179] concluded that UK general practice had removed many patients with depression from the denominator for QOF and switched instead to the use of symptom codes, which could also happen with OA coding.

3.6.5 Implications

A change in GP coding patterns to encourage the most specific diagnosis possible, such as required by the Morbidity Statistics in General Practice (in which GPs were asked to "[identify] *the problem(s) or diagnosis(es) which in the opinion of the general practitioner most comprehensively describes the morbidity for which the patient consults*" Fleming 1993 p.38[180]) would be of potentially substantial benefit in identifying a population of people who have clinical OA, and who may therefore be

candidates for structured disease management. At study inception, it appeared that there was too little specificity in some joint pain diagnoses to have complete confidence in the apparent population burden of OA as determined from general practice N05.. Osteoarthritis and allied disorders Read codes.

3.7 Template implementation study

3.7.1 Background

Although GPs and practice teams are well-used to templates as a routine part of their work, it was considered necessary to undertake some training in the use of the MOSAICS template to try to standardise the clinical use of the template for the purposes of the trial. The template was intended as a recording aid rather than a quality improvement intervention. The aim of the implementation strategies was to try to encourage its use as a recording tool for consultations considered possible to represent a diagnosis of clinical OA as noted above.

3.7.2 Method

The training was deliberately low-key, and emphasised the role of the template as an aide to consultation recording rather than suggesting it as a care pathway to be followed. To that end, the type of extensive training seen with the implementation of reminder systems (as reported in studies included in the Cochrane review[152]) was not used in this study. Clinicians were encouraged to diagnose and treat OA as they normally would, but to use the template to help structure their medical record. No link was made during the training between the template and the quality indicators from which the domains had been derived.

It was recognised that the use of the template added to practice's usual work. In order that this did not result in financial loss (and thereby reduce the likelihood of practice participation in the study), the study contract entered into between the practice and the research network included a payment for use of the template. Practices subsequently randomly allocated to the intervention arm of the trial had additional payments for the further work this entailed. The details of the financial arrangements were not part of the candidate's remit, but he advised that the payment be relatively

modest, and to be based on expected prevalence of OA and joint pain codes that would trigger the template rather than on the performance of the practice in using the template (to avoid as far as possible the risk of testing financial incentives alongside the template).

The template was introduced to the eight study practices in meetings with GPs and practice nurses, undertaken by the candidate or his clinical academic or Primary Care Research Network (PCRN) GP Research Facilitators (GPRF) colleagues, depending upon availability. A formal appointment was scheduled for this training, weeks or months in advance, arranged by the Keele study coordinator. The template was demonstrated using the live EMIS system and a (so-called) "dummy" patient, where available, and using a mock-up template on a stand-alone laptop in other cases. The single, approximately one hour, introductory meeting provided opportunity for the practice staff to ask questions. The candidate and the GPRFs emphasised the consultation recording aspect of the template (to de-emphasise the use of the template as a quality of care prompt) (Appendix E.7, page 390).

As well as the face-to-face training, the practices received electronic versions of the presentation slides for reference and a training DVD, both created by the candidate. The intention was that practices could use the DVD if required, as a refresher or to train new starters or clinicians who had been absent from the meeting.

The DVD was kept relatively brief and covered the mechanics of the triggering of the template, its completion, and the way in which the template and its data fitted into the medical record. Screen capture software was used to demonstrate the ways in which the template could be completed. No reference was made within the training materials to the quality indicators or the study more broadly. The script (which represents the notes made before the DVD creation rather than a verbatim transcript) upon which the DVD was based is reproduced in Appendix E.8.

After practice training, the template was installed by the informatics staff in each practice. The six-month pre-randomisation template phase of the MOSAICS study commenced at an agreed date

shortly after installation. The study commenced with the eight practices at two different time points, with four practices commencing at each point.

Further practice visits were undertaken by the MOSAICS study team, (the study Chief Investigator, the academic GPs including the candidate, and the GPRFs), three to four months after the introduction of the template, to remind practices to use it and provide an opportunity for feedback. These review meetings were once again of up to an hour's duration and there was an opportunity for the practice staff to ask questions. Comments from practices were captured on a brief data report form, with the following list of issues explored (data shared with a separate social science study[181,182]):

a. Do all GPs use the template (or only a proportion)?
b. For those GPs who use the template, how do they assess its usefulness with regard to diagnosing OA and treatment decisions?
c. For those GPs who do not use the template, what reasons do they give for not doing so?
d. Have the GPs talked to each other about the template? If so, do they have a shared perception [of the template feasibility and utility]?

As an interim method of assessment of the template, data were collected from participating practices after three to four months' template use. This assessment was originally referred to as an audit, though it does not have the qualities needed for such status, given the absence of established criteria and standards. The descriptive statistical analysis of these data was undertaken by the candidate's principal supervisor due to the requirements of the main study for immediate information. These data were presented to practices, with anonymous peer-referencing of achievement, at the template review meetings, similar to a light educational outreach visit.[183]

3.7.3 Results

Of the initial eight practice template training visits, the candidate attended five. For the template review visits, the candidate participated in five of the eight, the remainder being conducted by other study staff.

Template instruction meeting feedback

At the template instruction meetings (often held in conjunction with meetings about other aspects of the MOSAICS study, and all before the commencement of template use), feedback was collected from practices as far as possible (though without going to the extent of this being an established qualitative study).

Themes emerging from this feedback were:

1. Some practices had insightful questions about the study methodology (mainly relating to the effect of the template itself on recorded disease coding and care, and on the study analysis needed to identify this effect)
2. One practice in particular was very enthusiastic about the study in general and thought that the template itself would add value to the clinical activity
3. The organisation of the meetings was very variable. Some practices had a dedicated meeting for the purpose, in others it formed a component of a wider meeting. The facilities available were variable (internet access, clinical computer system access, projectors) and the proportion of staff attending was also very different between practices.
4. There was a general sense that the template was less applicable to the practice nurses, since they do not traditionally do much musculoskeletal clinical practice
5. It was not clear how the template training would be cascaded to other clinical staff in one practice
6. Some practices questioned the pain scoring system, identifying that they would have been happier with a numerical rating scale. Another practice questioned the need to record pain

and function at all, since patient-level outcomes were not to be assessed through this means

7. One practice had expected the template to be much more complex, such as to offer a direct (within-template) referral system to physiotherapy. Another practice had hoped to link the MOSAICS template directly with other service-driven templates (such as on falls and osteoporosis)

The overall impression of the training events was that the practices were willing to use the template but regarded it as something to accept for the purposes of the research rather than as something with intrinsic value to clinical practice. Two of the practices seemed to be keen to use the study to access additional benefits for the practice, staff, and patients, in terms of additional resources (financial, professional input from additional nursing time) and training.

Template review audit

Based on the prevalence figures in CiPCA, above, it was expected that the template would be triggered in the eight study practices 1900 times in the first four months (three months for one practice). The actual figure was 1470 times. However, it was also projected that there would be 1004 individual patients consulting with an OA trigger code recorded, whereas the actual figure was 1093 (158/10,000 registered population).

The template was used (i.e. had at least one entry made) on 828/1470 occasions (56%). The template was triggered but bypassed on the remaining 642/1470 (44%) of occasions. Information was entered on all eight indicators (excluding weight and BMI) on 378/1470 (26%) of occasions. There was variation in template use between practices, with a range of completion of at least one template item of 39%-79% of patients recorded as consulting with clinical OA; 9%-54% of patients consulting had all eight template items completed.

Table 3-6 shows the aggregated frequency that entries were made when the template had at least one entry. Pain and function assessment were template domains the most frequently populated, with weight advice the least frequently. These were not rates of quality indicator achievement

(discussed in Chapter Six) but rather the frequencies with which any data was entered into each of the template fields. To some extent, this may be regarded as demonstrating an order effect, since pain assessment appears at the top of the template and physiotherapy referral at the bottom.

Table 3-6: Frequency of template item completion where at least one entry made

Domain	**Patients with data entry *n* (%) (n=828)**
Pain assessment	776 (94)
Function assessment	746 (90)
Record of paracetamol use/advice	715 (86)
Record of topical NSAID use/advice	606 (73)
Information provision	602 (73)
Advice regarding weight	502 (61)
Advice regarding exercise	615 (74)
Advice regarding physiotherapy referral	575 (69)

Template review meeting feedback

Additional information collected from comments at the template review meetings was summarised as follows:

1. The repeated triggering of the template in the same patient over time was raised as an issue, with clinicians preferring the idea that the template would have stayed silent once completed, at least for a defined time period.
2. Self-reported estimates of 'escape' rates to avoid use of the template ranged up to from "not much" to 85% for one GP. The fact that the template triggered even when the clinician did not consider the diagnosis to be one of OA was brought up repeatedly. Another example was of a nurse practitioner using the template whenever it triggered whether or not the patient presentation was considered to be clinical OA.
3. Some GPs felt the template was over-inclusive for clinical OA, with it being a particular issue in younger patients. One practice would have preferred no age restriction as the GPs felt that it would have also applied to some younger patients.
4. Other staff felt the template was easy to use and had in fact become quite routine.
5. One practice stated that the template had no utility as an aid to diagnosis; another felt there had been no change to diagnostic practices as a result of the template.

6. Lack of flexibility in the response options was mentioned in two practices. A longer list of options would have been preferred, to include for example the use of low dose analgesics (paracetamol and topical NSAIDs).
7. Cascaded training for new staff was patchy, with some practices having only one member of staff present at the original training and no subsequent cascaded training despite the materials available.
8. Some practice staff had a less than positive view of the template – "a list of instructions," "generally a bit irritating," though the views expressed were those of individuals, as no practices seemed to have an uniform practice view of the template.
9. The implied recommendation for use of topical NSAIDs did not always mesh with clinicians' therapeutic preferences, though they felt generally that if the prompt was there because NICE had endorsed their use, it must be "OK" and so could be recommended.
10. There was a frustration in one practice that the template "doesn't go anywhere", "doesn't do anything," which, on further exploration, transpired to be a wish for it to actually constrain practice rather more than it did, and to generate referrals, such as to physiotherapy, automatically.

3.7.4 Discussion

The data collection at the three to four month point suggested that the template was working reasonably well, in terms of the frequency with which data was entered, and broadly in line with internal study team expectations. Although the audit assessment identified variability in consultation and completion rates, and the narrative feedback also highlighted a range of beliefs about and attitudes towards the template, it was decided to make no subsequent adjustments to the template. This was in part due to study time constraints, the fact that the study practices had already been exposed to the template, and the requirement to use the template as an aid to recording relevant information rather than as a quality improvement tool in its own right. The individual practice achievement was provided in face-to-face review meetings between the practices and the study team including academic GPs, and compared to the aggregated

achievement. This strategy of peer-referencing achievement was a recognised method of improving outcomes (in this case, the intended outcome being the template completion rate), although audit and feedback does not generally lead to large improvements in professional practice.[184]

Standardisation of the template response options was received with some criticism by the participating clinicians. It has previously been shown that alert systems open to customisation are associated with a better detection and resolution rate of prescribing problems[154] than fixed alert systems. It was unclear whether allowing practices to customise the OA template would have changed the participation rate but in any event the resulting heterogeneity of responses through allowing customised responses would have reduced the data utility for the clinical trial.

A method by which the training, as a study intervention, might be made more formal would have improved standardisation. This might take the form of a written prior agreement with practices to be available for standardised training to be used in conjunction with a detailed training schedule, including equipment lists, presentation outline, and expected attendance. It would have been helpful to have a third-party observer (a social scientist) present at all the meetings to identify the reactions of the recipients, and collate any themes arising in questioning across practices.

In terms of testing the template, it would have been preferable to be able to undertake a formal preliminary study with wider testing and resultant data analysis and qualitative feedback in practice, comparing template use to a control group, with the same Read codes available for recording information in each (rather than bespoke codes in the template group). Once so tested, the template could then have been implemented in the MOSAICS study practices to test the main intervention. However, the constraints imposed by various factors including practice availability, funding, and time meant that this approach was not possible.

3.7.5 Implications

The audited use of the template, although suboptimal, represented a big improvement on the routinely-recorded data for measuring quality achievement in OA care. There was considerable heterogeneity in the initial training and interim practice visit narrative feedback, also reflected in

variation in the use of the template as demonstrated by the audit assessment. Although it was not considered possible to adapt the template structure and trigger terms to address all of the feedback obtained, practitioners seemed to be content to maintain its use during the study period. Lack of adaptability for the template, due to the need for a constant approach across all practices, may have limited the degree of clinical engagement. Although this could not be overcome for a trial setting, such customisation could be delivered for implementation of the template in routine practice.

3.8 Conclusion

There is some evidence about the effects of reminder systems on recording and clinical practice but this effect seems small-to-moderate at most. This supports the view of the template predominantly as a consultation recording aid rather than as an intervention likely to cause a substantial change in clinical practice. There was little identified evidence to inform creation of the template. Although the quality indicators to be included were identified following a rigorous systematic review, it was not possible to define an optimum methodology for developing and testing a consultation recording template from previous literature. This left scope for the use of a pragmatic approach, in keeping with previous RIPCHS experience.

In terms of the requirements of the MOSAICS study, the template development, training delivered and implementation operated broadly within the expected parameters. When the template was used (at least one item completed), it was generally well-completed. Although coded data recording through the template identified in the audit phase was much better than identified previously through the routine record, template use may have been limited by clinical practice in that it was dependent upon the clinician recording data obtained in a consultation. This required collection of the information by the clinician in the first place, and then sufficient time and motivation to record the data in the template. Two factors in particular raised the possibility that data capture through the template in the MOSAICS study might not be as complete as desired: (i) the additional time estimated to be required for coding clinical information in an EHR (30-40 seconds[149]) and (ii) the

fact that template audit data showed substantial variation in template completion between practices (in the context of previous studies suggesting that motivation was important for reminder adherence[152]). Another option might have been to use patient supplementation of the EHR with patient-reported achievement of quality indicators. EMIS already have a facility for patients with some smartphones or similar devices to transmit information to a healthcare provider.[185] This does not yet facilitate writing of information directly to the EHR but rather to send messages to the healthcare provider or to create a personal health record. An extension of such technology to create a patient's own space to record information within a unified EHR would be another potential route toward collecting the information needed to track quality of care over time, as well as facilitating potential cross-validation of quality as recorded through the medical record.

Chapter Four will consider the baseline (pre-template) coding of OA and joint pain as a marker of existing quality of care, leading into the baseline recorded quality of care reported in Chapter Five and the quality of care captured through the template itself in Chapter Six. The effects of the template on recorded quality of care is examined in Chapter Seven.

Chapter Four: Patterns and associations with clinical OA coding in MOSAICS practices: a cross-sectional analysis

4.1 Introduction

Clinical coding of morbidity and processes of care in the UK is conducted primarily through the use of Read codes.[146] In a consultation, a GP might typically enter one or more Read codes to reflect a diagnosis or clinical problem, with associated coded assessment items (for example, weight and BMI measurements), uncoded narrative text, and (coded) drug treatment. Clinicians in general practice do not always make a formal diagnosis, sometimes choosing to record a symptom code.

The way in which UK primary care clinicians use Read codes has been found to be highly variable for conditions as diverse as osteoporosis,[186] gastrointestinal infections,[187,] and diabetes.[188] The general practice IM&T DES,[189] from 2006/07, went some way to ensuring a more consistent general level of disease recording and improved information management in UK primary care though it did not include any assessment of joint pain or OA diagnostic rates. Although there have been incentives for some specific coding behaviours (most notably to do with diagnoses of certain conditions and processes of care relating to the QOF,[102] and recently in dementia[190]), in general GPs are left to code clinical contacts in whatever way they prefer, and dependent upon their own beliefs and idiosyncrasies.[191]

A 2004 systematic review[192] of morbidity coding in general practice highlighted that, although it is possible to achieve high-quality coding, there are factors associated with the quality of morbidity coding in computerised general practice systems. These factors included

- the nature of the morbidity (the presence of "clear diagnostic features," such as numerical thresholds for diabetes, increased the likelihood of the diagnostic code appearing in the record). Although there is a definition of clinical osteoarthritis (Chapter One, section 1.2.2), this is vulnerable to clinical interpretation and diagnostic uncertainty

- the newness and perceived importance of the problem (with ongoing or potentially less important issues possibly being less well reported).[193] Given the fact that patients may consult about joint pain before eventually receiving an OA diagnosis,[194,195] and the reported negative attitudes (mainly by GPs) to OA in consultations,[63] OA risks not being seen as having sufficient 'newness' or 'importance' to be coded
- the "enthusiasm of practices and of individual GPs"[196]

Influences on diagnostic coding may occur at the locality, practice, clinician, or patient level. Localities may influence recording depending upon the style and extent of support and monitoring (for example, Read code use training, data quality facilitation, QOF assessments). At the level of the practice, factors may include the cultural coding norms and external use of the recorded patient data such as contribution to research databases such as CPRD. No empirical evidence was identified in the literature that evaluated these influences for OA. No evidence regarding the influences on OA coding behaviour at a clinician level were identified, though Bertakis et al.[197] reported various influences apparent on a primary care diagnosis of pain, including physician practice style. Other work in primary care has established an association between health status and physician practice style (better health being associated with a relatively low ratio of time spent history taking compared to examination).[198] Other characteristics such as the individual health beliefs of the practitioner (e.g. perceptions of OA as a process of aging, or as an obesity-related condition) and communication style between practitioners and patients could be important factors and in principle may interact with patient-level characteristics.[62,63,199] At a patient level, there are many potential factors that may influence coding behaviours, such as the presence of known risk factors for OA.[17]

The capture of data for research purposes is not typically one of the primary aims of a clinician entering data and in consequence the data quality for research purposes may be further reduced. Hersh et al.[171] identified potential limitations to high-quality coding, including inaccurate or incorrect data, the use of narrative data which cannot easily be used in research, and consultations about multiple problems that may appear under only one Read code with information about comorbidities in the narrative (or omitted altogether). These limitations risked some under-

ascertainment of cases of clinical OA in the MOSAICS study, though the highly sensitive definition of clinical OA was expected largely to have ameliorated this.

The use of a broad set of Read codes to define clinical OA was discussed in Chapter Three, section 3.5. One benefit of such an approach was to reduce the risk of selection bias, as clinicians could not easily avoid a diagnostic term and remove the patient thereby from the denominator population needing care for OA. However, with the adoption of such an approach, there are risks to the generalisability of the MOSAICS study. The use of a wide definition of clinical OA brings a greater risk of misclassification of people who do not have OA into the population regarded as having clinical OA. This could affect the external validity of the study, as only a proportion of the total study population actually having OA would be eligible for processes of care and so the proportions of people identified as having received appropriate care may be erroneously low.

Prior work to assess the utility of the medical record in determining the presence of OA in a patient (through a computerised algorithm[200] compared to a gold standard manual medical record data abstraction) has demonstrated limited sensitivity: 45%[201] to 78%[202] of OA cases were identifiable through the algorithm, though with a positive predictive value of 63%.[203] Harrold et al.[204] sampled an administrative database for people with and without a recorded encounter with OA in 1994-1996; of those with such an encounter, only 62% were rated as having definite OA from a medical record review with an additional 10% having possible OA, and of those without a recorded OA encounter, 18% were considered to have definite OA. In another sample of people with a recorded OA encounter from the same dataset, Harrold et al.[205] found a validated OA diagnosis in 63%. A previous case-control study of the prevalence and history of knee OA in one North Staffordshire practice identified an estimated 12.5% prevalence of knee OA ever recorded in people aged 45 years and over, but only 2.4% “currently recorded” (a recorded knee OA consultation within a 2-year period) , and the median time of initial knee symptoms to a diagnosis of knee OA was 10 years.[194] This serves to illustrate the complexity of primary care coding for OA and joint pain, with apparently the same underlying condition being coded as OA or as a pain-based symptom code, and reinforces the need for a wide Read code definition of a working diagnosis of OA.

The overall aim of this chapter is to investigate potential determinants of a recorded formal OA diagnosis rather than joint pain record. First, the extent of variation in the prevalence of OA or joint pain was assessed with respect to clinicians and practices. Secondly, the strength of any possible association of (i) patient-level risk factors for OA (age,[44 45] sex,[206-209] BMI[16,210-214], multimorbidity,[215] and X-ray information[12]) and (ii) potential clinician-level factors (volume of OA or joint pain consultations) with OA diagnosis rather than joint pain record was assessed. Finally, the extent of potential misclassification resulting from assuming that recorded joint pain in older adults is due to underlying OA was assessed through review of the narrative content of joint pain consultations and exploration of factors that may be used to discriminate between OA and 'not OA' will be investigated.

4.2 Methods

4.2.1 Comparison of recorded formal OA diagnosis and joint pain symptom codes

The reference period for usual coding practice in this analysis was taken to be the 12 months prior to template installation (baseline, phase one: see Figure 1-1, page 20). For practical purposes, the definition of the commencement of active involvement in the MOSAICS study was taken to be the installation and commencement of use of the template at the start of phase two. Although there were some prior meetings to describe the study (to gain participation) and set-up meetings, these were not considered to be sufficiently detailed to have caused a substantial change in practice within the phase one timeframe used for this analysis.

The index clinician and index consultation

Patients were allocated an *index clinician*. This was the first clinician to record a diagnosis of OA for that patient within the 12-month baseline period or, if no OA diagnosis was recorded, the first clinician to record a joint pain code in the period. The consultation with the first OA or joint pain record is the *index consultation*.

Case definition

Cases were defined as patients aged 45 years and over with a Read code within phase one of the MOSAICS study (12 months prior to template installation) of clinical OA (as defined in Chapter Three, section 3.5) affecting only peripheral joints (hand, hip, knee or foot). Although shoulder and elbow OA codes triggered the recording template, there were no corresponding joint pain trigger codes for these sites and so people consulting with only a record of the N05.. codes for these sites were removed from the numerator population for the analysis reported in this chapter. All other patients were allocated to the formally-diagnosed OA group if there was any recorded N05.. code in phase one (regardless of whether there was also a joint pain code), and to the joint pain group if there was a record only of the specified joint pain codes in phase one. Together, the formally-diagnosed OA and joint pain groups formed the overall clinical OA group.

Population of interest

The denominator population for the study was all patients aged 45 years or over registered at the point of the data collection (end of phase two, 2012). A static population for the MOSAICS practices was assumed.

Assessed characteristics and hypotheses for associations with an OA diagnostic code

The medical record data downloaded included all consultation records including reason for consultation, all prescriptions, and patient registration information. Also downloaded were relevant coded laboratory tests, radiology requests and results and referral data.

Various factors might have an influence on a clinician's tendency to use OA or joint pain codes. Based on the literature referred to in the introduction to this chapter, it was considered possible that the following recorded variables may have an effect:

i. Practice
ii. Clinician (hypothesis: that a large amount of residual variation in OA diagnostic coding would be at the level of the clinician)
iii. Number of (index) consultations each clinician undertakes, dichotomised into clinicians undertaking greater or fewer than the median number of index consultations

iv. Patient sex
v. Patient age (hypothesis: that older patients would be more likely to be given an OA diagnosis since age-specific rates for OA increase with age)
vi. Patient BMI status (hypothesis: that people who are overweight or obese would be more likely to be given an OA diagnosis since weight is a recognised risk factor for OA)
vii. Multiple patient consultations for OA/joint pain
viii. Site(s) of symptoms or diagnosis
ix. Patient morbidity burden
x. Use of X-rays (of hand, hip, knee, or foot) within phase one (see Figure 1-1, page 20) but before the index consultation (hypothesis: that people with a recorded X-ray would be more likely to be given an OA diagnosis since a radiology report identifying OA may prompt recording of a formal diagnosis)

Patients were allocated to one of four categories based on their most recent BMI prior to the index consultation in phase one as recorded in the dataset: BMI <25 (normal), 25-29.9 (overweight), ≥30 (obese), or unknown. All recorded heights and weights for patients consulting in phase one were available back to three years before template installation. This method assumed a steady state for BMI in a population with clinical OA. No empirical evidence applicable to such a population was identified to support or refute this assumption, though for older adults BMI tends to decrease with time; substantial changes occur over relatively long periods compared to the three years considered here.[216] Patients without a numeric BMI value but in whom a Read code of 'overweight' or 'obese' status had been recorded were allocated on the basis of the Read code type (e.g. C380. *Obesity* was allocated to the obese (BMI ≥30) category).

Measurement of total morbidity burden

The total burden of morbidity was determined through a count of drugs prescribed in the 12 months prior to the index consultation (the first OA consultation in a period or, if none, the first joint pain consultation), based on individual BNF[217] subchapters (the "BNF Chapter count"). This process has been found to be a simple proxy marker for prediction of future consultation and mortality[218] (second to the Charlson Index score[219]). Polypharmacy has sometimes been defined as the use of five or more drugs,[220,221] with more extreme polypharmacy defined as 10 or more.[220,222,223] Based on this approach, the number of BNF subchapters prescribed in the 12 months

before the index consultation was collapsed into ordinal categories of 0-4 drugs, 5-9 drugs, and 10 or more drugs. The use of subchapters reduces the risk of underestimating polypharmacy that may occur with higher-level aggregation of prescriptions as it would be unusual to take more than one drug from the most detailed subchapter classification (with the exception, for example, of emollients for eczema).

The Read codes used in phase one were mapped to anatomical sites as shown in Appendix E.9. Patients with recorded joint pain at more than one site during phase one were designated as having multi-site joint pain, and patients with recorded OA at more than one site (including generalised OA) were designated as having multisite OA.

Statistical analysis

Period consultation prevalence of OA and joint pain

Age-standardised rates (ASRs) allow the underlying rates of events in distinct populations with differing age structures to be compared without the potential masking effect of differences in age structure.[224] Using the direct method, the ASR for a particular condition is that which would have occurred if the observed age-specific rates for the condition had applied in a given standard population. Directly ASRs have been calculated using the methodology outlined by the ONS.[225] The European Standard Population (ESP) 2013 has been used.[226] This does not differentiate in age structure between males and females, contrary to the observed female excess in older age groups in the UK.[227] Age bands have been collapsed to 45-64, 65-74, 75-84 and 85+.

Crude and adjusted rates for OA and joint pain by practice are shown, as well by sex, age, site, BMI status, morbidity load, multiple clinical OA consultation status, and prior X-ray use. Results are presented as rates per 10,000 population.

OA diagnostic coding: associated factors

The outcomes of interest (the presence of an OA or joint pain code in the medical record) occur at an individual patient level. From clinical experience, it is well-recognised that individual doctors tend to attract a certain type of patient ("Doctors get the patients they deserve"[228]) and clinicians

also will have their own beliefs about OA and its coding. It was also possible that practices may function within themselves more consistently than between practices, due to, for example, practice-level clinical protocols. This had a potential clustering effect on the observed outcomes, which would affect the statistical analysis if not corrected for, such as within a multilevel model. The analyses of factors associated with receipt of an OA rather than joint pain diagnosis were assessed using multilevel binary logistic regression models (patients within index clinicians) for unadjusted and adjusted effects. The multilevel model was estimated using iterative generalised least squares with second-order penalised quasi-likelihood approximation (PQL2). Interaction effects between sex and age in the odds of an OA diagnosis were assessed due to the observation of a potential interaction in the descriptive epidemiology, and random slope effects for sex, age, and BMI status across clinicians were also assessed to determine whether or not their associations with OA diagnosis varied across clinicians. The data were also re-structured as a three-level model (patients within clinicians within practices) and the model re-run to determine whether there was additional practice variation in OA diagnosis. Results are presented as odds ratios (OR) with 95% CI.

4.2.2 Identification of potential OA misclassification

This aspect of the study was designed to identify which patients out of the population recorded as consulting with joint pain but not formally-diagnosed with OA (N05..) in phase one could be determined on the basis of narrative information from a joint pain consultation to have definite OA, 'not OA', or joint pain that could not further be classified from the information available. Following this, combination of the reclassified joint pain-only consulters with those formally diagnosed with OA, facilitated an estimation of the true annual consultation prevalence of OA in the MOSAICS practices.

For phase one, those patients identified as having consulted with joint pain and no recorded OA diagnosis in the same 12-month period were subject to a manual review of all joint pain consultations and their associated downloaded consultation narrative after export to a Microsoft® Access® 2013 database. Patients with joint pain but not an OA diagnosis who had been found to

have consulted with inflammatory joint disease prior to the index consultation (up to two years before the start of MOSAICS phase one) were removed from the dataset prior to the recoding exercise.

The basis for the recoding decision was (perforce) clinical judgement, but informed by the EULAR criteria for diagnosis of knee OA and the NICE guidelines on management of OA in adults.[5,11] The evidence relating to methods of diagnosis of OA is set out in Table 4-1.

Table 4-1: Guide for interpretation of narrative for reallocation of joint pain consultations

	Evidence for probable OA	**Evidence for 'not OA'**
History	"OA", gonarthrosis, "? OA" or variants Degenerate, degenerative, wear and tear, wear and repair, worn Persistent (>6 weeks) pain and/or stiffness Joint pain worse at end of day or with use, eased by rest	Red flags (*rapidly progressive symptoms, weight loss, morning stiffness > 30 minutes, identified **recent** significant trauma*) Other diagnosis substantially more likely (e.g. other apparent musculoskeletal diagnosis, neuropathy, arterial or venous disease, etc.) Can include people with symptoms e.g. of rapid or recent onset who do not meet other criteria for OA diagnosis.
Examination	OA, "?OA" Heberden's/Bouchard's nodes (H & B nodes), Other bony enlargement Painful, restricted movement at the hip, knee or foot Crepitus Changes in the joint suggestive of OA, including hallux valgus, other valgus or varus deformities	Red flags *(large effusion, redness or other signs of inflammation)*
Test results in text	Osteoarthritis, OA Loss of joint space Osteophytes, sclerosis, or cysts on plain X-ray MRI report suggesting OA Degenerate, degenerative signs	Red flags (e.g. *any radiology comment advising other pathology to be considered or referral made*)
Other opinion	Other clinical opinion cited (e.g. musculoskeletal or orthopædics) indicating OA	

Initially, two clinical investigators (including the candidate) independently (and blind to the other's decision) examined the consultation narrative for evidence that the patient could actually be regarded as having OA. The consultations linked to a joint pain code were assessed and allocated on the balance of clinical probability to OA/not OA/not further classified, as in Box 4-1.

Box 4-1: Reclassification categories for joint pain consultations

(i) Probable OA

for example, in records where OA is explicitly mentioned as the cause for the coded joint pain, or in cases which seem to be in keeping with the diagnostic criteria listed in Table 4-1.

(ii) Not OA

for example, records which identify a non-OA cause for the coded joint pain or in which the information given is divergent from the diagnostic criteria for OA (perhaps including inflammatory symptoms or signs, recent significant trauma, red flag symptoms) as described in Table 4-1.

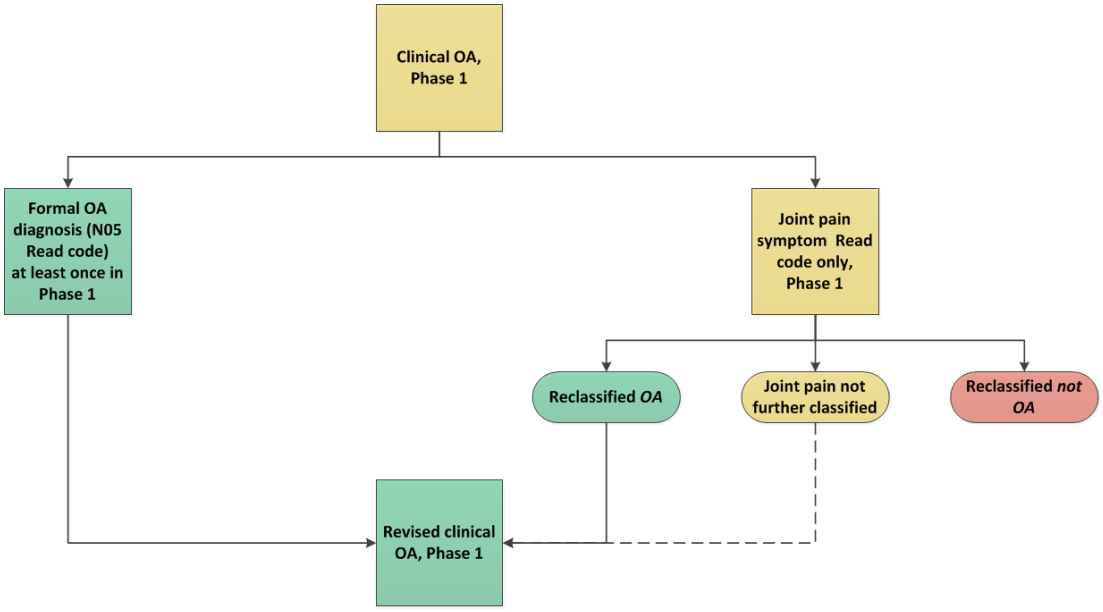
(iii) Joint pain not further classified

Insufficient information to permit classification: this category was used only for records which cannot sensibly be allocated to one of the other two.

Where no agreement was achieved between the two reviewers, a third GP clinical reviewer was asked to assess the consultation and independently provide an opinion about its classification as above. A simple majority decision was taken as binding, except where all three reviewers disagreed (one reviewer allocating the consultation to each of the three categories), in which case the consultation was determined to belong to the joint pain not further classifiable group.

Using all the consultations for clinical OA in phase one, patients with any joint pain consultation defined as probable OA were reclassified to the OA group; of the remainder, any 'not OA' consultation excluded the patient from the full clinical OA group. All remaining patients were left in a joint pain not further classified group, as shown in Figure 4-1).

Figure 4-1: Reclassification of patients with only joint pain codes



Revised estimates of consultation prevalence for clinical OA

After removal of all patients who had a recorded consultation for an inflammatory arthropathy in the two years prior to the start of phase one, three revised estimates of the consultation prevalence of clinical OA were derived: (i) patients with joint pain only who were reclassified as having OA were added to the formally-diagnosed (N05..) group and all other joint pain-only consulters were assumed not to have OA; (ii) those patients in the not further classified joint pain group were allocated to the clinical OA group on the basis that all were truly OA; (iii) those patients in the not further classified joint pain group were allocated to the OA and 'not OA' groups in the same age-specific ratios as those patients previously re-allocated to OA or 'not OA'. This final approach was considered to represent the best estimate of the valid consultation prevalence for OA in this population.

4.2.3 Data cleaning

The data obtained were cleaned by the CRN informatics team and Professor Jordan.

4.2.4 Statistical analysis

Analysis of joint pain misclassification

The numbers of consultations subjected to the reclassification exercise using the consultation narrative and the outcome (OA, joint pain not further classified, not OA) were identified and the agreement between raters measured by the (weighted) *kappa* statistic.

Revised estimates of consultation prevalence for clinical OA

Revised annual consultation prevalence estimates were produced using the methodology and three approaches described above. Results are presented as rates per 10,000 population.

OA, 'not OA', and joint pain: associated factors

For patients with only joint pain consultations in phase one, associations between the independent variables specified above and 'not OA' as an outcome of reclassification relative to reclassified OA or joint pain not further classified were estimated in a two level multinomial logistic regression model, to explore what variables were more or less strongly associated with the possibility of reclassification of joint pain to OA or 'not OA'. Random slope effects were tested. Results are presented as relative risk ratios (RRR) with 95% CI.

Variance partition coefficient

The variance partition coefficient (VPC) was estimated thus:

$$VPC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

where σ_u^2 is the level two residual (unexplained) variance and σ_e^2 is the level one residual variance, fixed at 3.29 in the logit model used. The resulting VPC, expressed as a decimal or percentage, estimates the amount of variance in the model accounted for by unobserved level two (clinician-level) variables. In a multilevel logistic regression model, the VPC needs to be interpreted with some caution since the fixed value for the level one residual variance means that there would be a scaling effect on the VPC when explanatory variables were added to the model.[229]

All analyses were undertaken in IBM SPSS v.21,[230] Microsoft® Excel® 2013,[231] and CIA software.[232] Multilevel models were conducted in Stata v.13[233] and MLwiN v.2.34[234] via runmlwin,[235] using PQL2, except where model convergence problems necessitate the use of other specified approximation methods.

4.3 Results

4.3.1 Denominator population structure

The denominator population totalled 33,726 (Table 4-2). One practice (ID number 1) was substantially larger than the other practices, accounting for more than one-third of the whole study population.

4.3.2 Morbidity coding of OA and joint pain

Consultation prevalence rates by practice

Annual consultation prevalence rates for OA and joint pain during phase one (crude and age standardised) by practice are shown in Table 4-3. Overall, the estimated consultation prevalence for combined clinical OA was 940 per 10,000. There was inter-practice variability with rates for joint pain varying between 404 and 839 per 10,000 and for OA varying between 83 and 495 per 10,000. Broadly, practices with higher rates of OA diagnoses had lower rates for recorded joint pain in people aged 45 years and over.

Table 4-2: Denominator population age and sex frequencies, by practice.

	Practice ID																	
	1		2		3		4*		5		6		7		8		Total	
Age band	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**	**Female (%)**	**Male (%)**
<45	6222	6337	1740	1987	1869	1912	1778	1773	1935	1896	863	1016	1835	1895	1030	1080	17272	17896
	(51.4)	(53.9)	(62.8)	(65.4)	(44.8)	(47.8)	(40.9)	(43.1)	(52.2)	(52.5)	(45.4)	(48.9)	(50.4)	(53.1)	(50.3)	(53.3)	(49.8)	(52.3)
45-64	3443	3423	632	680	1294	1223	1302	1260	1011	1021	544	650	1006	1011	586	581	9818	9849
	(28.4)	(29.1)	(22.8)	(22.4)	(31.0)	(30.6)	(29.9)	(30.6)	(27.3)	(28.3)	(28.6)	(31.3)	(27.7)	(28.3)	(28.6)	(28.6)	(28.3)	(28.8)
65-74	1214	1142	214	215	514	520	624	600	423	432	261	261	424	410	222	221	3896	3801
	(10.0)	(9.7)	(7.7)	(7.1)	(12.3)	13.0)	(14.4)	(14.6)	(11.4)	(12.0)	(13.7)	(12..6)	(11.7)	(11.5)	(10.8)	(10.9)	(11.2)	(11.1)
75-84	830	646	136	122	339	275	438	369	240	219	153	118	260	212	156	126	2552	2087
	(6.9)	(5.5)	(4.9)	(4.0)	(8.1)	(6.9)	(10.1)	(9.0)	(6.5)	(6.1)	(8.1)	(5.7)	(7.1)	(5.9)	(7.6)	(6.2)	(7.4)	(6.1)
85+	399	212	48	36	152	72	206	111	101	46	79	33	113	40	55	20	1153	570
	(3.3)	(1.8)	(1.7)	(1.2)	(3.6)	(1.8)	(4.7)	(2.7)	(2.7)	(1.3)	(4.2)	(1.6)	(1.6)	(1.1)	(2.7)	(1.0)	(3.3)	(1.7)
Total all ages	12108	11760	2770	3040	4168	4002	4348	4113	3710	3614	1900	2078	3638	3568	2049	2028	34691	34203
Total aged ≥45 years	5886	5423	1030	1053	2299	2090	2570	2340	1775	1718	1037	1062	1803	1673	1019	948	17419	16307
	(48.6)	(46.1)	(37.2)	(34.6)	(55.2)	(52.2)	(59.1)	(56.9)	(47.8)	(47.5)	(54.6)	(51.1)	(49.6)	(46.9)	(49.7)	(46.7)	(50.2)	(47.7)

*One female in practice 4 is missing due to a lack of birth date information in the dataset but is not included in the OA or joint pain consultation population and would form a part of the denominator only.

Table 4-3: Age-specific and directly age-standardised annual consultation prevalence rates by practice, peripheral joint pain or OA in phase one

		Age-specific and standardised recorded disease prevalence rates per 10,000 (95% CI) by practice								
	Age band	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Combined**
Joint pain	**45-64**	465	442	600	574	748	528	724	386	549
		(415,519)	(336,571)	(508,704)	(485,674)	(634,877)	(405,675)	(611,851)	(281,516)	(516,582)
	65-74	573	443	716	629	971	728	839	474	670
		(480,678)	(267,692)	(562,898)	(496,786)	(773,1203)	(515,999)	(654,1060)	(293,725)	(614,731)
	75-84	589	426	1010	843	1089	701	911	355	754
		(472,727)	(213,763)	(774,1294)	(654,1068)	(809,1436)	(422,1095)	(659,1227)	(170,652)	(677,838)
	85+	589	476	670	915	612	1250	850	400	714
		(413,816)	(130,1219)	(375,1104)	(613,1314)	(280, 1162)	(683,2097)	(452,1453)	(82,1169)	(593,852)
	All aged 45+	510	442	688	654	842	638	783	402	613
		(469,554)	(356,542)	(613,770)	(584,729)	(748,944)	(535,756)	(692,881)	(318,501)	(587,640)
	Age-standardised rate* per 10,000	518	438	674	644	839	628	787	404	617
		(475,560)	(348,527)	(598,750)	(574,715)	(743,935)	(522,735)	(693,880)	(315,493)	(590,643)
OA	**45-64**	281	228	167	164	212	42	109	163	200
		(243,324)	(154,326)	(120,226)	(118,222)	(153,285)	(14,98)	(68,165)	(98,254)	(181,221)
	65-74	649	350	319	335	480	153	156	474	424
		(551,761)	(196,577)	(220,448)	(240,454)	(344,651)	(66,302)	(83,267)	(293,725)	(379,472)
	75-84	854	426	277	335	588	148	191	957	530
		(711,1017)	(213,763)	(161,443)	(220, 487)	(388,856)	(40,378)	(87,362)	(631,1393)	(466,601)
	85+	1064	357	357	536	272	0	654	267	612
		(821,1356)	(74,1044)	(154,704)	(312,859)	(74,697)	(0,329)	(313,1202)	(32,963)	(509,751)
	All aged 45+	475	283	228	259	329	81	144	351	318
		(436,517)	(216,365)	(185,277)	(216,308)	(272,395)	(47,130)	(107,190)	(273,444)	(299,338)
	Age-standardised rate* per 10,000	495	303	220	247	330	83	148	354	323
		(453,537)	(226,380)	(177,263)	(204,290)	(270,390)	(43,122)	(107,189)	(270,437)	(304,342)

* European Standard Population, age 45 years and over

Table 4-4: Consultation prevalence for peripheral joint pain or OA in phase one with aggregated crude and directly standardised period consultation rates

Crude annual consultation prevalence rates	Joint pain			OA			Clinical OA		
	Female	Male	F:M ratio	Female	Male	F:M ratio	Female	Male	Both sexes
45-64	627 (579,679)	471 (429,516)	1.3	260 (229,294)	142 (120,168)	1.8	887 (829,948)	613 (565,664)	750 (712,789)
65-74	757 (673,849)	587 (512,669)	1.3	526 (457,603)	318 (264,380)	1.7	1283 (1173,1401)	905 (812,1006)	1097 (1024,1173)
75-84	886 (774,1009)	594 (494,709)	1.5	615 (523,719)	426 (342,525)	1.4	1501 (1354,1659)	1021 (888,1167)	1285 (1184,1392)
85+	703 (558,873)	737 (531,996)	1.0	598 (466,757)	667 (472,915)	0.9	1301 (1101,1527)	1404 (1113,1747)	1335 (1168,1519)
All aged ≥45 years, per 10,000 (95% CI)	699 (661,740)	523 (489,559)	1.3	394 (365,424)	238 (215,263)	1.7	1168 (1116,1221)	761 (719,805)	933 (900,966)
Age-standardised* annual consultation prevalence rate per 10,000 (95% CI)	703 (664,743)	527 (492,563)	1.3	391 (361,420)	252 (227,278)	1.6	1094 (1045,1143)	780 (737,823)	940 (907,973)

* European Standard Population, age 45 years and over

Clinician effects

There were 115 index clinicians identified, with a median of 10 patients (range 1-129, interquartile range (IQR) 2-51). Of 60 clinicians holding the median or fewer index consultations, 32 coded no consultations as OA and eight coded all as OA, whereas only one of the clinicians holding more than the median number of index consultations coded none as OA, though none coded all consultations as OA.

Patient-level variable effects

Age and sex effects: age-specific and standardised consultation prevalence rates

The age-specific prevalence for both joint pain and OA was highest in women in the 75-84 years age band, and in men in the 85+ band (Table 4-4). The higher prevalence for females was less evident in the older age bands.

Site of disease: consultation prevalence rates

Table 4-5 gives the annual consultation prevalence rates by age and site. The site with the highest consultation prevalence was the knee for both joint pain (age standardised rate 317/10,000) and OA (age standardised rate 135/10,000).

Body mass index

As shown in Table 4-6, BMI status analysis shows that there was an increasing number of patients consulting in each of the diagnostic categories as the BMI increased, though the category with the most persons consulting was 'BMI unknown,' at around ⅓ of people consulting. An OA diagnosis relative to joint pain was most frequent (38.1%) in the obese (BMI ≥30kgm^{-2}) category, consistent with known risk factors.

Morbidity burden

The ratio of OA to joint pain diagnoses increased with increasing morbidity, with approximately twice the ratio in the highest morbidity band compared to the lowest.

Prior X-ray use

People with recorded use of an X-ray in phase one before the index consultation for OA or joint pain received a diagnosis of OA compared to joint pain more frequently than those in whom no X-ray was recorded (ratio of 0.45 OA:joint pain in those with no X-ray increasing more than five-fold to 2.45 where a prior X-ray was recorded (Table 4-6).

Multiple consultations

Patients with more than one consultation for clinical OA recorded in phase one were recorded as having a diagnosis of OA rather than joint pain more frequently than those with only a single recorded consultation (Table 4-6).

Table 4-5: Age-specific and directly age-standardised annual consultation prevalence rates in phase one by site, diagnosed peripheral joint pain or OA

Age-specific rates per 10,000	Knee		Hip		Ankle/Foot		Wrist/Hand		Unspecified		Multiple	
	Joint pain	OA	Joint pain	OA	Joint pain	OA	Joint pain	OA	Joint pain	OA	Joint pain	OA
45-64	304	80	101	36	45	9	58	18	14	17	27	41
65-74	344	169	165	88	60	8	66	23	18	45	19	90
75-84	334	237	250	80	73	9	60	24	11	45	26	136
85+	302	290	279	70	41	6	70	17	6	70	17	168
All aged 45 years and over	317 (298,337)	133 (121,145)	145 (133,159)	56 (48,64)	52 (45,60)	9 (6,12)	61 (53,70)	20 (15,25)	14 (10,19)	30 (25,37)	25 (20,31)	71 (63,81)
ASR per 10,000*	317 (298,336)	135 (123,148)	146 (133,159)	57 (48,65)	53 (46,61)	9 (6,12)	61 (53,69)	20 (15,24)	14 (10,19)	31 (25,37)	25 (20,30)	72 (63,81)

*Age-standardised consultation prevalence, European Standard Population, age 45 years and over

Table 4-6: Characteristics of patients consulting for peripheral joint pain or OA in phase one

		OA *n* (%)	Joint pain *n* (%)	OA:JP ratio	Total *n*
	Total	1074 (34.1%)	2071 (65.9%)	0.52	3145
Site	**Knee**	447 (29.5%)	1069 (70.5%)	0.42	1516
	Hip	188 (27.7%)	490 (72.3%)	0.38	678
	Ankle/foot	29 (14.2%)	175 (85.8%)	0.17	204
	Wrist/hand	67 (24.6%)	205 (75.4%)	0.33	272
	Unspecified	102 (68%)	48 (32%)	2.13	150
	Multiple	241 (74.2%)	84 (25.8%)	2.87	325
BMI status	**BMI <25**	141 (31.5%)	307 (68.5%)	0.46	448
	BMI 25 to <30	252 (33.8%)	493 (66.2%)	0.51	745
	BMI 30+	346 (38.1%)	561 (61.9%)	0.62	907
	Unknown	335 (32.1%)	710 (67.9%)	0.47	1045
BNF chapter count	**0-4**	456 (29.0%)	1118 (71.0%)	0.41	1574
	5-9	331 (36.6%)	574 (63.4%)	0.58	905
	10+	287 (43.1%)	379 (56.9%)	0.76	666
Consultations in phase 1	**Single**	536 (28.2%)	1368 (71.8%)	0.39	1904
	Multiple	538 (43.4%)	703 (56.6%)	0.77	1241
Relevant XR recorded in phase 1, prior to index consultation	**No**	910 (31.2%)	2004 (68.8%)	0.45	2914
	Yes	164 (71.0%)	67 (29.0%)	2.45	231

Multilevel logistic regression analysis

The associations with a diagnosis of OA compared to joint pain only, unadjusted and adjusted for each of the factors, are shown in Table 4-7.

There was generally only a small difference between the results from unadjusted and adjusted models. Factors with a significant association with a recorded OA diagnosis were patients with older age: adjusted OR compared to age 45-64 of 1.83 (95% CI 1.47,2.27) for age 65-74 rising to 2.29 (1.61,3.26) for age 85+; patents with unspecified (adjusted OR 7.32, 95% CI 4.62,11.0) or multisite disease (adjusted OR 6.07, 95% CI 4.44,8.30) compared to knee; patients with obesity compared to not overweight (adjusted OR 1.42, 95% CI 1.06,1.90); patients who had multiple consultations for

clinical OA compared to single (adjusted OR 1.43, 95% CI 1.19,1.73), and recorded prior use of X-rays compared to not (adjusted OR 4.89, 95% CI 3.46,6.91). Those with ankle/foot disease were less likely to have an OA diagnosis compared to knee (adjusted OR 0.44, 95% CI 0.28,0.69).

The VPC was estimated in the null (no independent variables) model at 0.19 (0.12, 0.25) and in the adjusted mode at 0.10 (0.05, 0.14) (that is, 10% of the residual variation in use of an OA code was due to unobserved clinician-level factors).

Due to an observed potential interaction effect seen in the descriptive epidemiology, the multilevel model was repeated with the sex*age interaction effects specified in the independent variables and the resulting ORs are also shown in Table 4-7. The point estimate for the OR for each of the interaction terms suggested a positive interaction, though it was only statistically significant for the male*age band 75-84 age band (OR 1.67, 95% CI 1.02,2.72), indicating that older age had a greater impact on increased likelihood of OA diagnosis in males.

Table 4-7: Associations of independent variables with an OA diagnosis in phase one

	Unadjusted OR (95% CI)	Adjusted OR (95% CI)	Interaction model adjusted OR (95% CI)
Sex (*reference: female*)	**↓ 0.80 (0.67,0.94)**	0.87 (0.73,1.04)	**↓ 0.74 (0.56,0.97)**
Age 65-74 (*reference: 45-64*)	**↑ 1.81 (1.49,2.20)**	**↑ 1.83 (1.47,2.27)**	**↑ 1.76 (1.34,2.31)**
Age 75-84	**↑ 1.98 (1.60,2.46)**	**↑ 1.93 (1.50,2.48)**	**↑ 1.60 (1.17,2.18)**
Age 85+	**↑ 2.27 (1.66,3.09)**	**↑ 2.29 (1.61,3.26)**	**↑ 1.98 (1.28,3.05)**
Male*Age 65-74 interaction	-	-	1.12 (0.72,1.73)
Male*Age75-84 interaction	-	-	**↑ 1.67 (1.02,2.72)**
Male*Age 85+ interaction	-	-	1.50 (0.74,3.04)
Hip (*reference: knee*)	0.98 (0.79,1.21)	0.85 (0.68,1.07)	0.85 (0.68,1.07)
Ankle/foot	**↓ 0.40 (0.26,0.62)**	**↓ 0.44 (0.28,0.69)**	**↓ 0.45 (0.29,0.70)**
Wrist/hand	0.80 (0.58,1.09)	0.97 (0.70,1.35)	0.97 (0.70,1.35)
Unspecified	**↑ 6.60 (4.43,9.83)**	**↑ 7.32 (4.86,11.0)**	**↑ 7.26 (4.82,10.9)**
Multisite	**↑ 7.01 (5.21,9.43)**	**↑ 6.07 (4.44,8.30)**	**↑ 6.08 (4.45,8.32)**
BMI 25 to <30 (*reference: BMI <25)*	1.09 (0.83,1.43)	1.12 (0.83,1.50)	1.12 (0.83,1.51)
BMI 30+	**↑ 1.42 (1.10,1.85)**	**↑ 1.42 (1.06,1.90)**	**↑ 1.44 (1.08,1.93)**
BMI unknown	1.03 (0.80,1.34)	1.14 (0.85,1.51)	1.16 (0.87,1.54)
5-9 BNF chapters (reference: 0-4)	**↑ 1.53 (1.26,1.84)**	1.17 (0.95,1.46)	1.17 (0.94,1.45)
10+ BNF chapters	**↑ 1.93 (1.57,2.37)**	1.23 (0.96,1.57)	1.23 (0.97,1.58)
Multiple consultations in phase 1 (*vs. single*)	**↑ 2.08 (1.76,2.45)**	**↑ 1.43 (1.19,1.73)**	**↑ 1.44 (1.19,1.73)**
X-ray in phase 1 prior to index consultation (*vs. none*)	**↑ 4.66 (3.37,6.43)**	**↑ 4.89 (3.46,6.91)**	**↑ 4.91 (3.48,6.93)**
Above the clinician median index consultation count (*reference: at or below the median*)	**↑ 1.79 (1.08,2.98)**	1.50 (0.91,2.47)	1.48 (0.90,2.45)
Practice 2 (*reference: Practice 1*)	0.70 (0.28,1.79)	0.91 (0.37,2.22)	0.90 (0.37,2.21)
Practice 3	**↓ 0.29 (0.16,0.51)**	**↓ 0.34 (0.20,0.60)**	**↓ 0.34 (0.20,0.60)**
Practice 4	**↓ 0.47 (0.26,0.86)**	**↓ 0.45 (0.25,0.80)**	**↓ 0.45 (0.25,0.80)**
Practice 5	0.56 (0.30,1.03)	0.66 (0.36,1.20)	0.66 (0.36,1.20)
Practice 6	**↓ 0.24 (0.09,0.61)**	**↓ 0.26 (0.10,0.67)**	**↓ 0.26 (0.10,0.68)**
Practice 7	**↓ 0.28 (0.15,0.50)**	**↓ 0.27 (0.15,0.50)**	**↓ 0.27 (0.15,0.50)**
Practice 8	0.94 (0.46,1.92)	1.10 (0.53,2.32)	1.09 (0.52,2.30)

Two-level multilevel logistic regression model, patients within clinicians

No significant random slope effect was identified across clinicians for sex effects. There was a small degree of evidence that the association between the odds of an OA code and both age and BMI varied slightly across clinicians.

A three-level model of patients within clinicians within practices showed no significant explanatory effect of the practice on the residual variance.

4.3.3 OA misclassification and revised consultation prevalence for clinical OA

There were 2058 patients coded with only joint pain in phase one, who had a total of 3209 consultations with a range of 1 to 16 consultations per patient, median one consultation and IQR 1-2 consultations.

The recoding exercise using the consultation text with the first two GP raters resulted in agreement in 2196 (68.4%) of the 3209 recoded consultations and a Cohen's *kappa* value for inter-rater agreement of 0.49, p<0.001, which indicated moderate agreement.[236] Introduction of a third GP rater resulted in a *kappa* value for multiple raters of 0.46 (moderate agreement) from a majority decision in 895 (27.9%) additional consultations and complete disagreement in 118 (3.7%) consultations.

Of the 2058 patients, 515 (25.0%) were recoded to 'truly OA', 718 (34.9%) to 'not OA', and 825 (40.1%) were left in the joint pain (not further classifiable) group.

The flow of patient numbers from the original clinical OA group to the three possible revised clinical OA groups is shown in Figure 4-2. If the 825 patients with unclassifiable joint pain consultations were assumed to belong to the OA or 'not OA' groups in the same ratio, stratified by age-group, as those which could be classified, this would add a further 483 people to the total misclassified (i.e. the 'not OA' group), to a total of 1201. Deduction of these from the 3126 people previously allocated to the clinical OA group (all OA plus defined peripheral joint pain), would equate to a crude annual consultation prevalence for 'true' clinical OA of 576 per 10,000 (95% CI 550,602). Allowing for the fact that it is uncertain what proportion of unclassifiable patients have OA, the potential range of crude consultation prevalence for true OA was 474 per 10,000 (where all unclassified joint pain patients were assumed not to have OA) to 720 per 10,000 (where all unclassified joint pain patients were assumed to have OA).

The age and sex category-specific annual consultation prevalence sensitivity analysis estimates for the revised clinical OA and 'not OA' populations are given in Table 4-8.

Figure 4-2: Derivation of revised clinical OA group in phase one

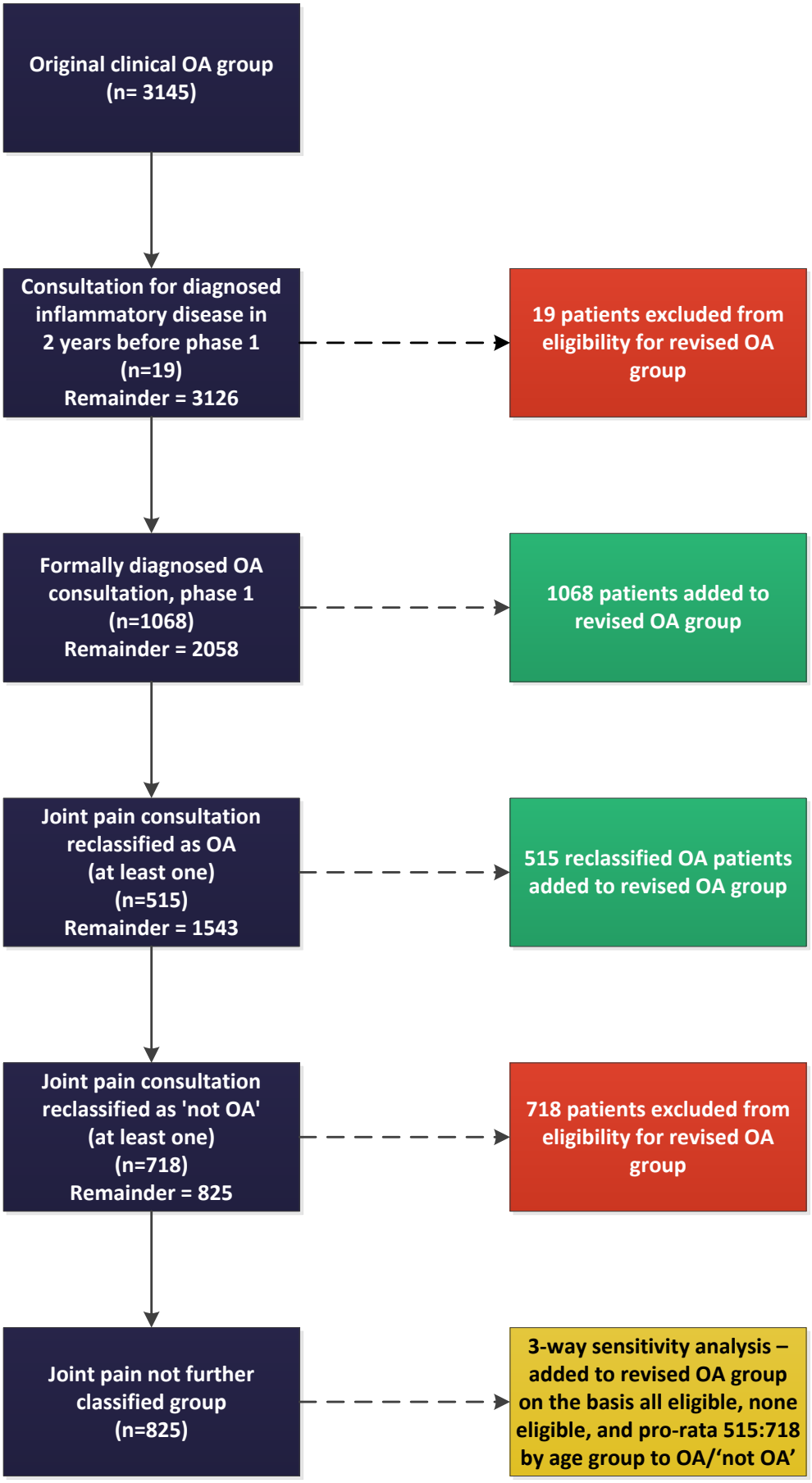


Table 4-8: Original and revised estimates of age-specific and directly age-standardised annual consultation prevalence rates per 10,000 for clinical OA

		Sensitivity analysis for final clinical OA group		
Age category-specific rates per 10,000	**Original clinical OA group, phase 1**	**Point (best) estimate for all clinical OA**	**Minimum possible clinical OA**	**Maximum possible clinical OA**
45-64	750	416	327	327
65-74	1097	712	599	854
75-84	1285	880	742	1043
85+	1335	875	784	1021
All aged 45 years and over	933	571	469	714
Directly age-standardised* rate per 10,000 (95% CI)	940 (907,973)	576 (550,602)	474 (451,498)	720 (691,749)

*European Standard Population, age 45 years and over
Original and revised estimates including a sensitivity analysis comparing predicted OA consultation prevalence rates after reclassification of patients with only non OA (N05..) joint pain consultations

4.3.4 Factors associated with joint pain misclassification

Only a small proportion (12.1%) of people with ankle/foot disease had definite OA in those recorded with joint pain and many (66.1%) had 'not OA' as shown in Table 4-9.

Table 4-9: Final patient classification (OA, 'not OA', or joint pain unclassified), by disease site

Site	OA *n* (%)	'Not OA' *n* (%)	Joint pain unclassified *n* (%)
Knee	304 (28.5)	322 (30.2)	440 (41.3)
Hip	113 (23.3)	141 (29.1)	231 (47.6)
Ankle/foot	21 (12.1)	115 (66.1)	38 (21.8)
Wrist/hand	39 (19.3)	88 (43.6)	75 (37.1)
Unspecified	3 (6.4)	20 (42.6)	24 (51.1)
Multiple	35 (41.7)	32 (38.1)	17 (20.2)

Associations with reclassification as true OA, 'not OA' or joint pain unclassifiable) are shown in Table 4-10.

Factors associated with reclassification of joint pain as true OA rather than 'not OA' were patients in age band 75-84 compared age 45-64 years (RRR 1.44, 95% CI 1.06,1.95), and patients with an overweight status (RRR 1.57, 95% CI 1.11,2.24) and obesity (RRR 1.45, 95% CI 1.03,2.05) compared to not overweight. Patients with knee disease had an increased risk of an OA diagnosis:

Table 4-10: RRRs for OA/not OA and joint pain/not OA, reclassified joint pain patients

	Clinical OA defined by joint pain codes only: comparison of reclassified groups	
	OA vs. 'not OA' comparison RRR[a] (95% CI)	**Unclassifiable joint pain vs. 'not OA' comparison RRR[a] (95% CI)**
Sex (*reference: female*)	0.81 (0.66,1.01)	0.90 (0.74,1.08)
Age 65-74 (*reference: 45-64*)	1.30 (1.00,1.68)	0.94 (0.75,1.18)
Age 75-84	↑ **1.44 (1.06,1.95)**	1.07 (0.81,1.41)
Age 85+	1.04 (0.65,1.67)	↓ **0.64 (0.41,0.97)**
Hip (*reference: knee*)	0.78 (0.60,1.02)	1.20 (0.95,1.50)
Ankle/foot	↓ **0.18 (0.11,0.29)**	↓ **0.22 (0.15,0.33)**
Wrist/hand	↓ **0.42 (0.28,0.62)**	↓ **0.55 (0.40,0.76)**
Unspecified	↓ **0.20 (0.06,0.67)**	1.12 (0.60,2.09)
Multisite	0.99 (0.61,1.62)	0.61 (0.34,1.07)
BMI 25 to <30 (*reference: BMI <25*)	↑ **1.57 (1.11,2.24)**	↑ **1.42 (1.05,1.93)**
BMI 30+	↑ **1.45 (1.03,2.05)**	1.34 (0.99,1.81)
BMI unknown	1.21 (0.87,1.70)	1.21 (0.90,1.62)
5-9 BNF chapters (*reference: 0-4*)	1.21 (0.94,1.57)	1.08 (0.85,1.35)
10+ BNF chapters	1.05 (0.77,1.43)	1.18 (0.90,1.55)
Multiple consultations in phase 1 (vs. single)	1.22 (0.98,1.52)	↓ **0.52 (0.42,0.64)**
X-ray in phase 1 prior to index consultation (*vs. none*)	0.88 (0.49,1.60)	0.97 (0.57,1.64)
Above the median for clinician index consultation count (*reference: at or below the median*)	1.21 (0.80,1.85)	↑ **1.95 (1.32,2.88)**
Practice ID 2 (*reference: Practice 1*)	1.37 (0.67,2.79)	↑ **1.80 (1.03,3.14)**
Practice ID 3	0.76 (0.50,1.16)	↓ **0.67 (0.47,0.94)**
Practice ID 4	↓ **0.47 (0.29,0.75)**	↓ **0.54 (0.38,0.77)**
Practice ID 5	0.86 (0.54,1.39)	0.74 (0.51,1.08)
Practice ID 6	0.57 (0.29,1.12)	0.96 (0.58,1.59)
Practice ID 7	0.93 (0.59,1.46)	↓ **0.62 (0.42,0.90)**
Practice ID 8	0.62 (0.30,1.28)	1.04 (0.59,1.83)

[a]Multinomial two-level logistic regression (patients within clinicians) relative risk ratios, adjusted for all covariates.

compared to distal limb or unspecified sites (with reference to the knee, risks of an OA diagnostic outcome were: ankle/foot RRR 0.18, 95% CI 0.11,0.29; wrist/hand RRR 0.42, 95% CI 0.28,0.62; unspecified site RRR 0.20, 95% CI 0.06,0.67; hip and multisite disease were not statistically significantly different to the knee).

For joint pain unclassified compared to 'not OA', factors associated with an increased risk were patients with an overweight status compared to not overweight (RRR 1.42, 95% CI 1.05,1.93) and an index consultation with a clinician who had more than the median number of index consultations (RRR 1.95, 95% CI 1.32,2.88). Patients in age band 85+ had a lower risk of a joint pain compared to 'not OA' compared to age 45-64 (RRR for joint pain compared to 'not OA' 0.64, 95% CI 0.41,0.97). Patients with a knee site of disease had an increased risk of a 'not OA' diagnosis than distal limb sites (compared to knee, risk of a joint pain outcome were: ankle/foot RRR 0.22, 95% CI 0.15,0.33; wrist/hand RRR 0.55, 95% CI 0.40,0.76; other sites were not statistically different to the knee). Multiple consultations were associated with reduced risk of a joint pain outcome (RRR for joint pain 0.52, 95% CI 0.42,0.64).

No random slope effects were identified for sex, age, or BMI status in the two-level models. Three-level models of patients within clinicians within practices demonstrated no statistically significant explained variance at the practice level and only small differences in the RRR estimates compared to the two-level model.

4.4 Discussion

4.4.1 Principal findings

Clinical OA and the use of OA or joint pain diagnoses

The age-standardised annual consultation prevalences for joint pain and OA were 617 and 323 per 10,000 respectively, with a combined prevalence for the clinical OA category of 940 per 10,000. There was extensive variation by practice for both recorded joint pain and OA diagnoses. Once misclassification of patients with only joint pain codes had been corrected for, the possible range for true clinical OA lay between directly age-standardised rates of 474 (95% CI 451,498) and 720 (691,749) per 10,000 per annum, with a best estimate of 576 (550,602) per 10,000.

The patient-level factors independently associated with increased odds of an OA diagnosis were increasing age, site (unspecified or multisite disease), obesity, multiple consultations for clinical OA,

and recorded X-ray use prior to the index consultation. Ankle/foot site of disease was associated with reduced odds of an OA diagnosis.

The association of prior X-ray use with an OA diagnosis suggested that clinicians may be using the results of X-rays to record a diagnosis of OA. This was not consistent with the NICE guidelines, which do not recommend routine use of radiological investigations to make a diagnosis of OA.[5]

The association of multiple consultations for clinical OA with a formal (N05..) diagnosis of OA may reflect an increased severity of symptoms requiring more frequent consultation, or a clinical review of undifferentiated joint pain leading to a diagnosis of OA subsequently (akin to surveillance bias[237]). Although multiple consultation refers to only consultations for clinical OA, another explanation might be clinical comorbidity; given the known association of OA with obesity (and therefore by extension to other conditions also associated with obesity that may also require consultation) and with increased mortality,[238-240] patients may potentially have primarily consulted about another matter and the clinical OA been recorded as a secondary issue. Clinical comorbidity with OA has previously been described as extensive, with propensity to consult as one identified possible explanation.[241] This analysis did not identify any link between overall morbidity burden as measured by BNF chapter count and OA diagnostic code in an adjusted multilevel model, however, though any association may have been confounded by the effect of age as both prevalence of OA and polypharmacy increase with age.[242]

There was no significant association between the numbers of index consultations (unique patients) undertaken by clinicians and an OA diagnosis, suggesting that musculoskeletal clinical workload did not affect coding behaviour.

Joint pain misclassification

The assessment of joint pain misclassification identified a substantial number (35% of those recorded as joint pain rather than diagnosed with OA) of patients who were included in the initial definition of 'clinical OA' who were determined at case note review to have consulted with joint pain that did not represent OA. This resulted in a lower point estimate of the annual consultation

prevalence for clinical OA in phase one (576 per 10,000 compared to the original estimate of 940 per 10,000) though there was uncertainty about the correct classification of people whose joint pain consultations could not reliably be allocated to OA or 'not OA'. The difference in the proportions of patients with joint pain reclassified to OA or 'not OA' (25% to 35%) suggests that there may be a difference in the type of narrative information recorded for people with joint pain according to the underlying working diagnosis. No other such exercise has been identified in the published literature and so this is a potentially important contribution to the understanding of the definition of clinical OA within health care databases.

Nearly two-thirds of patients with an ankle/foot joint pain code were reclassified as 'not OA', and 43.6% of those with wrist/hand disease. Taken with the association between ankle/foot and wrist/hand sites of disease and increased adjusted risk of a 'not OA' outcome after reclassification, this indicates a need for caution in attribution of patients with a joint pain coded consultation to a clinical OA group. Joint pain codes with an unspecified site of disease were infrequently reclassified as true OA and care regarding their inclusion in a clinical OA definition should be exercised.

The associations between a diagnostic OA code and a joint pain symptom code showed some consistency with the associations in the reclassified joint pain OA compared to 'not OA'. Age and overweight/obesity were common associations. Site effects showed some differences between the original and reclassified comparisons, though this may represent coding artefact due to the use of N05.. OA diagnostic codes that either do not specify a site (such as N05.. Osteoarthritis and allied disorders) or are known to be multisite (such as N050. Generalised osteoarthritis - OA). X-ray result information was not routinely available to the clinical reviewers of the joint pain consultations, unless referred to in the consultation narrative, which may explain the lack of association between prior X-ray use and a reclassified OA diagnosis in the assessment of misclassification. Multiple consultation was associated with an OA diagnosis compared to joint pain but in the reclassification assessment was associated only with increased risk of 'not OA' compared to joint pain not further classified. This may reflect a process of refinement of a diagnosis, either formally coded or reflected in the narrative, over more than one consultation.

Agreement between the first two clinicians, including the candidate, was moderate at best, but the solution of introducing a third clinical reviewer still left considerable difficulty in achieving agreement on the reclassification decisions. This suggests that the narrative information linked to the joint pain consultations was limited in its description of the working diagnosis.

4.4.2 Comparison with existing literature and implications

This analysis has demonstrated that recording of an OA diagnosis was less common than a joint pain code and was associated with risk factors for development of OA such as age[243-245] and obesity.[246-249] The findings were not dissimilar to the analysis by Jordan et al. which linked (for those patients consenting) self-reported information to medical records in responders to a survey.[17] That study was developed subsequently to complement and assess the generalisability of the analysis reported here and to allow inclusion of self-reported information, with the candidate inputting into the design and interpretation and a co-author on the publication. That study also found substantial variation between practices in coding an OA diagnosis or joint pain, and association of age and obesity with an OA diagnosis, as well as self-reported pain interference. The analysis reported here used the entire practice population consulting with clinical OA, rather than restricting analysis to patients who both responded to a survey and consented to medical record review. This analysis also included a measure of morbidity load (a count of prescription drug types) and use of X-rays, assessment of interactions between patient characteristics and between patient characteristics and clinicians in a multilevel model, and novel interrogation of narrative text.

It has not hitherto been clear to what extent the proportion of people coded as having joint pain actually have OA. Interrogation of narrative text to establish likelihood of true OA in those recorded with joint pain leading to identification of an estimated rate of true clinical OA that is greater than that found by use of an OA diagnosis alone (but less than combined prevalence based on the highly-sensitive set of joint pain or OA codes) is novel. Comparison of diagnosed OA consultation prevalence in different UK primary care datasets demonstrated a substantial spread in rates with 3.6-19.2% of all musculoskeletal consultations recorded as OA, and absolute rates of 38 to 426 per

10,000 people aged 15 and over: considered to be most likely due to differences in recording practices.[21] International comparisons of OA prevalence show only moderate differences between UK primary care consultation prevalence rates (aged 45 and over) and Swedish at 375 and 443 cases per 10,000 people respectively for OA and 1074 and 967 per 10,000 respectively for OA or joint pain.[23] The rates obtained in this analysis were a little lower than generally reported at 284 per 10,000 for OA, and a combined OA or joint pain annual consultation prevalence rate of 940 per 10,000. This may in part reflect the use of the European Standard Population for age-standardisation. However, unlike the CiPCA database[22] used in the UK-Swedish comparison, the MOSAICS practices had not been through a prior regime of training and data quality assessment which may partly explain the difference.

There remains a need to develop a standard definition of OA in administrative and primary care clinical datasets, as was highlighted by Harrold et al. for administrative data in 2000.[204] This would not only assist the use of EHR databases for research purposes but would also facilitate a structured approach to care through use of recall systems and quality improvement tools such as audit.[100] This has to some extent been tackled for come conditions through the QOF[102] for conditions like COPD and asthma, which have to meet diagnostic criteria, and heart failure, which requires referral for echocardiography or a specialist opinion within a set period after diagnosis. This type of disease validation is harder to achieve for OA especially given the working diagnosis mandated by the National Collaborating Centre for Chronic Diseases[4] as set out in Chapter One, section 1.2. That is to say, there is no external gold standard by which a diagnosis can be made.

Delivery of enhanced diagnostic coding of OA in primary care would probably require a combination of strategies. A previous assessment of coding has concluded that current systems promote diversity rather than consistency in coding behaviours[250] and this may reflect a need further to develop the use of diagnostic coding in practice. There is evidence that coding can be improved by a programme of training and feedback.[22] Tai et al.[250] critiqued existing methods of code selection in general practice and recommended a more limited set of diagnostic codes to limit unwarranted

variation in coding. Overall, enhanced morbidity recording seems to be feasible as demonstrated by the differences between databases noted above.[21] Better morbidity coding might be regarded both as a marker of better quality of care in its own right and as a necessary prerequisite to delivery of better care for OA in general practice.

4.4.3 Strengths and limitations

This analysis has a number of strengths. The use of an entire population of general practice population consulting with clinical OA has enabled an assessment that is free of the selection bias associated with some analyses, and should be more generalisable to general practice in the UK as a whole. The range of independent variables, which included known OA risk factors as well as adjustment for morbidity levels and a measure of a clinician's OA workload (as the index consultation count), has reduced the risk of confounding of the associations.

Practice 1 formed a large part of the total study population, which may have caused the results of the investigation to be unduly weighted by the behaviours of this practice and the characteristics of its practice population.

Although the variables of interest were decided upon before the analytical work, there were multiple comparisons made within this analysis with no adjustment of the accepted level of statistical significance to reduce the probability of a type 1 error. However, Rothman argued that correction for multiple comparisons was not to be undertaken without a consideration of a corresponding rise in the risk of a type 2 error.[251,252] Since interpretation has focussed more on consistent themes rather than on individual comparisons, an appropriate balance between these two types of error seems likely to have been struck.

Many patients did not have a BMI record captured by the data download and some BMI records were not contemporaneous with the code used to allocate people to an OA group or joint pain. The apparent effect of BMI on coding behaviour may have been modified by the missing and non-contemporaneous data. Documentation of weight status in the medical record may be more common with increasing BMI.[253] To assess the effect of this potential bias, a sensitivity analysis was

conducted (not shown), reclassifying of all the BMI unknown patients to BMI <25 kgm^{-2}, but this had very little effect on the adjusted ORs.

The method of allocating patients to a morbidity load based on BNF chapter counts is a relatively new approach.[218] The categories used of 0-4, 5-9 and 10+ drugs were based on prior work that used drug counts used in a 3-month period rather than the prior year.[222] If an OA code represents one end of a severity spectrum (as inferred from the findings of Hensor et al.[254] that knee OA is preceded by knee joint pain), one might expect that more drug groups (analgesics) be prescribed to people with an OA code. The use of a BNF chapter count as a measure of total morbidity may therefore cause the apparent morbidity level to be slightly inflated as the analgesics prescribed for OA add modestly to the apparent total burden. Classification of people to the morbidity groups of 0-4, 5-9 and 10+ on the basis of 12 months' prescriptions may not be generalisable to other definitions of polypharmacy that use only a week's prescriptions in the assessment; it is however consistent with the work by Brilleman et al.[218] Other morbidity measurements such as the Charlson Index as adapted for UK primary care codes[255] would be an appropriate comparator methodology to validate the effect of morbidity itself.

This analysis has examined associations with an OA diagnosis within a 12-month period, including an assessment of repeated consultation by way of a dichotomy between single and multiple clinical OA consultations. Ideally frequency of repeated consultations would be considered, though few patients in this study had more than two consultations for clinical OA. If clinicians 'follow suit' in coding behaviour, it is also possible that coding in subsequent consultations would have been influenced by the initial consultation (i.e. the codes used were not independent of each other at each time point). The potential for considerable complexity, especially where there is a crossover in coding between OA and joint pain within a period, made a repeated-measures approach unfeasible for this analysis. At the level of the patient, the effect of individual deprivation on coding behaviour would have been interesting to investigate: there is some evidence of an effect of social class, occupation, or deprivation on the incidence of musculoskeletal pain,[256-260] though the effect on coding behaviour is unknown. A history of relevant joint trauma would also be relevant to

examine for effect on coding, given the evidence for trauma in the causation of some OA.[212,249] Although only 10% of the residual variance in coding OA and joint pain was estimated to be explained by unobserved clinician characteristics, factors that would have been desirable to analyse include those identified by Clarson et al. as significantly associated with agreement for an OA domain for the QOF or with monitoring OA (special interest in musculoskeletal disease, higher research or Master's degree, familiarity with the NICE OA guidance, and working full-time). At the practice level, neighbourhood-level deprivation has not previously been identified to be significantly associated with OA incidence[261] and so an effect on coding is less plausible. Practice size was not found by Broadbent et al.[69] to be associated with quality achievement though it is possible that this or other practice-level variables, such as being active in education, research, or commissioning, may be associated with coding behaviour.

Initially, a decision rule for the reclassification of joint pain to OA or 'not OA' was trialled but the information contained within the available consultation data was insufficient to apply a rule set. Despite several attempts to formalise the process, and to record the reasons behind the conclusion reached for each consultation, it remained a very subjective process (albeit one undertaken by three experienced clinicians). The process of agreeing on a reclassification outcome was limited as there was no measure of intra-rater reliability due to lack of availability of clinician time.

4.4.4 Conclusions

Currently-used estimates of OA prevalence based on administrative datasets or primary care EHRs seem likely to be inaccurate as those that exclude all joint pain are insufficiently sensitive but those that accept joint pain as a synonym for OA are insufficiently specific. On the basis of the data presented in this analysis, a definition of clinical OA from the primary care EHR may be considered to be more specific and valid if it excluded joint pain codes for sites of ankle/foot and wrist/hand.

Factors associated with increased odds of an OA diagnosis rather than joint pain code were mainly those known to be associated with increased risk of OA. It is not known to what extent the presence

of these risk factors was influencing a clinical decision to code a patient as OA rather than joint pain even where the underlying problem was similar.

Development of improved coding techniques for OA and joint pain in primary care is needed better to define OA for ongoing clinical management and for OA studies in clinical databases. Analysis of joint pain coding to assess which codes were more likely to represent OA in the population as a whole or within subpopulations may help to make a definition of OA based on GP-recorded Read codes a viable prospect with the potential for better opportunity for structured disease management.

Having considered here the clinical OA morbidity coding in phase one or MOSAICS, the next chapter will consider the quality of care for OA in general practice as measured from routinely-recorded data in the same phase.

Chapter Five: Baseline quality of care achievement for OA: a cross-sectional analysis

5.1 Introduction

This chapter will provide an overview of the routinely recorded assessment and management of clinical OA in primary care. It is intended to show how the routinely coded electronic record can be used to assess some aspects of the recorded quality of care for OA and to highlight its limitations in measuring quality of care. It acts as a prelude to the quality assessment made through the template, which is discussed in the next chapter.

Previous assessments of OA care have been described in Chapter One: success criteria used previously to determine the quality of care for OA have been variable. Although the OA quality indicators proposed in Chapter Two have an apparently clear numerator and denominator, the implementation of these or similar indicators in previous studies has tended to be dependent upon a somewhat broader interpretation of achievement of the quality criterion. Li et al.,[80] in an assessment of the non-pharmacological care of OA through a survey self-report, used 'severe' or 'extreme' pain on walking as the denominator and at least one visit to a physiotherapist or occupational therapist in the preceding year as the numerator (success criterion) for the quality indicator relating to assessment for ambulatory assistive devices for people with difficulty walking for at least three months. Similar interpretations were used for non-ambulatory assistive devices. Broadbent et al.[69] used well-defined indicators but it was not apparent from the published article exactly how the numerator was arrived at – for example, clinical experience would suggest it unlikely that routine records would contain details that patients with OA had been offered information about its natural history, treatment, and self-management and yet there was a 30% pass rate identified for this indicator; a wider interpretation of the success criteria than that suggested by the indicator is therefore presumed. Influences relevant to UK general practice on recorded quality of care for OA have been suggested to include: patient age (different effects by

indicator), sex (female associated with higher achievement than male) and disease severity (severe OA associated with better achievement than less severe).[69]

As noted in Chapter Three, there are many aspects of care for OA that are not the subject of routine coding, including the assessment of pain and function, information and education provision, and assessments of over-the-counter drug use. Zingmond et al.[79] also found a lack of routine administrative data (similar to the electronic, coded information used in the MOSAICS medical record review study) hampered use of various indicators such as referral to an exercise programme, use of paracetamol, and oral NSAID risk communication. The assessment of quality of usual care described in this chapter is dependent upon those aspects of care that were routinely recorded and this required a similarly inclusive approach to the definition of success criteria as described above.

The aims of this analysis were (i) to describe the baseline level of recommended and non-recommended processes of care for OA in general practice identifiable from routinely recorded information (prior to the template installation), (ii) to validate which quality indicators for OA in general practice were usable without modification to recording practices, and (iii) to estimate associations between quality achievement and characteristics of patients and clinicians.

5.2 Method

5.2.1 Denominator population

This analysis was conducted on the same group of patients included in the analysis reported in Chapter Four, with the same use of an index consultation and index clinician as described in section 4.2.

5.2.2 Assessed characteristics

The independent variables assessed for their association with recorded achievement in this analysis were the same as those described in Chapter Four, section 4.2.

5.2.3 Quality indicators and their success measures

All of the processes analysed in this chapter are linked to quality of care as described by the NICE OA guidelines[5] or the systematic review of primary care OA quality indicators described in Chapter Two.[107] An additional indicator, relating to non-use of oral NSAIDs in the presence of relative contraindications, was derived by the candidate from advice in the British National Formulary (BNF).[217] The indicators are listed in Table 5-1.

Table 5-1: Quality assessment indicator domains for routinely recorded management of OA

Indicator	Source
'To do' indicators	
Assessment of pain	Systematic review
Assessment of function	Systematic review
Evidence of education or advice for OA	Systematic review
Exercise or physiotherapy referral	Systematic review
Exercise assessment	Systematic review
Weight record	[Prerequisite for weight loss advice indicator]
Weight loss advice for people with a working diagnosis of OA and a BMI ≥25kgm^{-2}	Systematic review
Assessment or referral for assistive devices (ambulatory or nonambulatory) – intended for people with functional impairment	Systematic review
Prescription for paracetamol	Systematic review
Prescription for topical NSAIDs	NICE guidelines[5]
Prescription for standard oral NSAIDs	Systematic review
Evidence of documentation of NSAID risk assessment (for people receiving a prescription for oral NSAIDs)	Systematic review
Co-prescription of gastroprotection (for people prescribed oral NSAIDs)	Systematic review
Prescription for capsaicin	NICE guidelines[5]
Prescription for orlistat (weight loss agent)	NICE guidelines[5]
Completion of all the above before specialist referral.	Systematic review
'Do not do' indicators	
Use of X-rays	NICE guidelines[5]
Prescription for oral NSAIDs in the context of relative contraindications	British National Formulary[217]
Prescription for etoricoxib 60mg (NSAID)	NICE guidelines[5]
Prescription for topical rubefacients	NICE guidelines[5]
Prescription for glucosamine	NICE guidelines[5]

The indicators used were dependent upon coded data which could be analysed straightforwardly in standard statistical software programs rather than on the narrative contained within electronic or paper records. Some indicators from the systematic review were not relevant to this assessment of care – primary prevention of OA through weight loss for people who are overweight was excluded as the denominator population was people with established clinical OA, and first-use of paracetamol could not reliably be established due to the time-limited nature of the dataset.

Individual and population-level indicators

For most indicators, an assessment of the quality of care depends on aggregation of recorded outcomes to a population level. Some indicators were applicable to all patients – for example, pain and function assessment should be conducted annually at least (see the proposed indicators resulting from the systematic review, Chapter Two, section 2.3.2). Weight loss advice for people known to be overweight would also be applicable every year. Other indicators would not necessarily be recorded for all patients each year and so should be regarded as population-level measures. The relatively safe pharmacological management options may not be appropriate on an individual level but at a population level, better care would be likely to be associated with higher levels of recorded prescription. Some aspects of care were considered in the systematic review[107] to be necessary once after an OA diagnosis (education, exercise prescription), or annually (exercise review, weight loss advice for people who were considered overweight). Some outcomes were considered to be relatively undesirable at a population level on the basis of the 2008 NICE OA guidelines.[5] In general, one might expect that the first-line agents (paracetamol and topical NSAIDs), along with gastroprotection for people prescribed oral NSAIDs, should be relatively commonly recorded compared to other agents. X-ray use should be relatively sparing given the NICE guidance on X-ray use.

All patients consulting with clinical OA should ideally be recorded as having received an assessment of pain and function, an exercise review, and weight loss advice where the patient is known to be overweight. For other indicators, it was not possible to set target levels for the proportion of

patients expected to be recorded as 'achieving' them (prescriptions, X-rays, referrals) as requirements for treatment would vary with case mix.[262]

Time period for quality indicator success

The QOF[102] uses a period of 12 months for achievement in many processes-of-care (such as asthma or diabetes reviews, and stroke risk assessment in atrial fibrillation) and six months for some prescribing processes (such as the definition of "currently treated" for some drugs after myocardial infarction or in heart failure).[263] A single instance of the qualifying code is typically sufficient under the QOF processes for an indicator to have been achieved, except for value-based data (such as blood pressure measurement or laboratory results), which require the latest entry before QOF year-end to be at or below a certain threshold. It was not possible to use the same approach for OA, as there were very few OA-specific process codes identified in the Read dictionary. Codes that could be used to define quality indicator achievement in OA were equally applicable to some other conditions. Although optimal temporal linkage would require the same date for processes of care and the clinical OA consultation code entry, a 14 days 'grace period' was added on the basis that at a consultation, the need for further assessment and treatment may be identified. Fourteen days was selected arbitrarily as the cut-off for such follow up, in line with previous practice at the RIPCHS for consideration of allocation of drug treatment to a consultation.[264] The potential for mis-attribution of care processes is acknowledged but it would be unusual for patients to consult general practice twice within a 14-day period as the median number of consultations per person per year was estimated at 5.4 (2008 figures).[265] It was also possible that some care processes (such as for repeat prescriptions, for example) may be related to clinical OA but fall outside the 14 day period. As a compromise between over- and under-inclusiveness, it was determined that quality indicators could be achieved by a record of the success criteria within 14 days of any consultation for clinical OA in phase one.

As a sensitivity analysis, the assessment codes, prescribing and referral data were also examined across the whole of phase one such that anyone consulting with clinical OA who had received the

process of care or prescription any time within phase one were included in the numerator for indicator achievement.

Morbidity assessment and advice indicators

The same Read codes were used as for the estimation of routinely-recorded care in CiPCA (Chapter Three, section 3.4). For each care process (assessment of pain and function, education, exercise assessment, exercise advice, weight loss advice, NSAID risk assessment), the indicator was regarded as achieved if there was a relevant Read code present in the electronic health record (EHR) within the time frame of interest. Records of relevant X-rays were treated in the same way.

Prescription indicators

The prescription of paracetamol, topical NSAIDS, opioids, oral NSAIDs, gastroprotection, capsaicin, and weight loss agents were assessed. It was not possible to identify from the records available what treatment, if any, had been trialled previously due to the time-limited nature of the data collection. Assumptions about quality of care were based on population levels of prescription, on the basis that relatively safe options (paracetamol, topical NSAIDs) should be prescribed more commonly than other analgesics within populations although at an individual level other treatments may be more clinically appropriate. Opioids in OA treatment have been identified in Cochrane reviews to have only small-to-moderate benefits and these were deemed to be outweighed by a substantial increase in risk of harm or withdrawal due to lack of tolerability.[266,267] A relatively low level of opioid prescribing would therefore be expected. Oral NSAID use could not be benchmarked against a gold standard target. NICE recommended that paracetamol and topical NSAIDs should be considered ahead of oral NSAIDs,[5] so one might anticipate a lower level of oral NSAIDs compared to the two relatively safe analgesic options.[267]

NICE mandate the use of PPIs for gastroprotection[5] and so patients prescribed oral NSAIDs in the period of interest were assessed for PPI prescription in the same period. Since not all patients tolerate PPIs well, an additional descriptive analysis examined prescription of other gastroprotective agents in patients not prescribed a PPI. For the analysis of oral NSAID prescriptions

within 14-days of a clinical OA consultation, achievement was taken as a PPI prescription within the same time frame. The sensitivity analysis examined all gastroprotection prescriptions within phase one.

The only licensed weight loss drug available for prescription during the study period was orlistat. This has a stated indication for use in people with a body mass index (BMI) of 30 kgm^{-2} or more, and in individuals with a BMI of 28 kgm^{-2}or more in the presence of other clinical risk factors.[217] This analysis examined only people with the higher (obesity) BMI threshold as the presence of additional clinical risk factors had not been determined. It was not possible to set an expected level of orlistat prescribing given the known problems with tolerability and effectiveness.[268]

Treatments recommended by NICE not to be given (etoricoxib 60mg, topical rubefacients, and glucosamine)[5] were also assessed, with an expectation of non-use or only low level use of these drugs (low-level use might be expected due to difficulties in negotiating treatment cessation if patients have been taking these agents with a perception of benefit previously, and sometimes clinicians do not adhere to treatment guidelines for other potentially valid reasons).

Relative contraindications to oral NSAID treatment

There are many potential absolute and relative contraindications to treatment with oral NSAIDs set out in the BNF.[217] Relevant comorbidities, identified on the basis of clinical experience augmented by the BNF[217] and Read code browser,[148] are shown in Box 5-1. This is not a comprehensive list of contraindications: others exist but were not selected due to limitations in retrieving some information from the electronically coded record (for example, NSAID-triggered asthma is not included). The use of a 24-month period was estimated (on the basis of an examination by the candidate of data from CiPCA[21,22]) to provide 90% sensitivity to the presence of a relevant co-morbid condition in the medical record compared to a 36-month period and so, for reasons of data download practicalities, a 24-month period prior to commencement of phase one was used to capture comorbidity.

It was not possible to set an expected level of oral NSAID prescription in the presence of recorded relative comorbid contraindications, as the clinical justification for such a prescription would vary depending upon OA severity, response to other treatments, comorbidity severity, and patient and clinician beliefs about balances of risks and benefits. The analysis was exploratory, with an expectation that levels of oral NSAID prescribing would be low where a relative comorbidity contraindication was present in the EHR within the two years prior.

Box 5-1: Comorbidities as relative contraindications to oral NSAIDs

Co-morbidity	**Read codes and terms**
- 195..	Indigestion symptoms (+ child codes)
- 1Z1..	Chronic renal impairment (+ child codes)
- G2…	Hypertensive disease (+ child codes)
- G3…	Ischaemic heart disease (+ child codes)
- G58..	Heart failure (+ child codes)
- G6…	Cerebrovascular disease (+ child codes)
- G70..	Atherosclerosis (+ child codes)
- G73..	Other peripheral vascular disease (+ child codes)
- J1...	Oesophageal, stomach and duodenal diseases (+ child codes)
- J60..	Acute and subacute liver necrosis (+ child codes)
- J61..	Cirrhosis and chronic liver disease (+ child codes)
- J622.	Hepatic coma
- J623.	Portal hypertension
- J624.	Hepatorenal syndrome
- J625.	[X] Hepatic failure
- J62y.	Other sequelae of chronic liver disease
- J62z.	Liver abscess and chronic liver disease causing sequelae NOS
- J63..	Other liver disorders (+ child codes)
- J68..	Gastrointestinal haemorrhage (+ child codes)
- K0...	Nephritis, nephrosis and nephrotic syndrome (+ child codes)
- R071.	[D]Heartburn. (+ child codes)

Referral quality indicators

Referral, within 14 days of a clinical OA consultation, to other services for exercise (physiotherapy or other specified exercise programmes), physiotherapy alone (as a proxy for ambulatory assistive devices), occupational therapy referral (as a proxy for non-ambulatory assistive devices), and specialist input was determined for all patients consulting in phase one. Referral at any point within phase one was also determined as a sensitivity analysis, though this was not used in the assessment of associations between referral outcome and the independent variables of interest.

5.2.4 Data analysis

Quality of care was described in terms of the percentage of patients achieving the indicators in phase one (baseline, 12 months prior to template installation), overall and stratified by the characteristics listed in Chapter Four (section 4.2), with the addition of diagnostic group (OA or joint pain at the index consultation) as an independent variable.

Multilevel logistic regression analyses (patients within index clinician, adjusting for practice through the use of dummy variables) were used to estimate associations between the baseline patient characteristics and the clinician index consultation count, and quality indicator achievement. Except where stated, models were estimated using second-order penalised quasi-likelihood (PQL2) approximations. Results are presented as ORs with 95% CI. The random slope effects of age, sex, and BMI by clinician were assessed, on the basis that studies have previously identified age and sex discrimination,[269-271] and clinicians have previously been found to be prone to negative stereotypes about obesity.[272] Adjusted three-level models, of patients within index clinician within practice, were tested as sensitivity analyses and are reported in Appendix F (the findings were similar to the adjusted two-level models). Some processes of care were too infrequently provided for any logistic regression model to be used successfully; these are identified and descriptive epidemiology used to assess patterns of provision.

The data were analysed in SPSS v21,[230] Stata 13.1[233] and MLwiN[273] using the runMLwiN[235] command. Results are presented as counts with percentages for the descriptive epidemiology, and OR with 95% CI for the multilevel models. Statistically significant associations in the adjusted two-level logistic regression models are shown in the results with positive associations in **blue bold** text and negative in **red bold**.

5.3 Results

Except where specified, in the multilevel model analyses no statistically significant sex*age interaction, or random slope effects across clinicians for sex, age, or BMI status was identified. In

the three-level models of patients within clinicians within practices no statistically significant explanation of variance at the practice level was identified.

5.3.1 Baseline characteristics

The baseline characteristics for the population of clinical OA consulters in phase one are shown in Table 5-2 (these were shown for OA and joint pain consulters in Chapter Four, Table 4-6). As shown in Chapter Four, the ratio of people recorded only with joint pain to those diagnosed with OA formally was approximately 2:1 (65.9% joint pain only). A majority were female (60.5%). Knee was the site of disease most frequently recorded (48.2%) followed by hip (21.6%) with ankle/foot (6.5%) and unspecified sites (4.8%) the least frequently recorded. Only 14.2% were known not to be overweight, 28.8% were known to be obese, and one-third had no recorded BMI status. Half of consulters were in the lowest morbidity band of 0-4 BNF chapters and 21.2% in the greatest. Few patients had an X-ray recorded within phase one before the index consultation (7.3%). Single consultations were recorded in 60.5% and multiple in 39.5% of consulters.

The denominator population for the majority of assessments was the population consulting with clinical OA in phase one (n=3145). For measurements of gastroprotection and etoricoxib 60mg, the denominator was the proportion of patients prescribed an oral NSAID (n=565). For the oral NSAIDs in the presence of a relative comorbidity contraindication, the denominator was the number of clinical OA consulters with a recorded comorbidity in the period of up to two years before the start of phase one (n=1599). Assessment of prescribing of orlistat (for weight loss) had a denominator population of 907 people (all those known to have a BMI≥30kgm^{-2}).

5.3.2 Quality indicator assessment

Patient assessment and advice indicators: Read-coded elements of consultations

No patients were recorded as receiving education or an oral NSAID risk assessment within 14 days of a clinical OA consultation or within phase one overall; these processes of care are not shown in Table 5-3.

Pain assessment was not recorded within 14 days of a clinical OA consultation in phase one at all (Table 5-3), and only six OA-consulters (0.2%) had a recorded pain assessment at any time within phase one overall. Function assessment likewise was not recorded within 14 days of a clinical OA consultation and only 11 patients (0.3%) had such an assessment recorded at any time in phase one. Exercise assessment was not recorded within 14 days of a clinical OA consultation but was recorded in nine (0.3%) clinical OA consulters within phase one overall. Exercise advice (three patients, 0.1%) and weight advice (one patient, 0.1%) were infrequently recorded, always within 14 days of a clinical OA consultation.

Table 5-2: Baseline characteristics of MOSAICS clinical OA consulters, phase one

		***n* (%)**
Total		3145
Diagnostic group	Joint pain	2071 (65.9)
	OA	1074 (34.1)
Sex	Female	1904 (60.5)
Age band	45-64	1477 (47.0)
	65-74	842 (26.8)
	75-84	598 (19.0)
	85+	228 (7.2)
Site of disease	Knee	1516 (48.2)
	Hip	678 (21.6)
	Ankle/foot	204 (6.5)
	Wrist/hand	272 (8.6)
	Unspecified	150 (4.8)
	Multiple	325 (10.3)
BMI status	Not overweight	448 (14.2)
	25<=BMI<30	745 (23.7)
	BMI>=30	907 (28.8)
	Unknown	1045 (33.2)
Morbidity load (BNF chapter count)	0-4	1574 (50.0)
	5-9	905 (28.8)
	10+	666 (21.2)
X-ray, phase 1[a]		231 (7.3)
Multiple consultations		1241 (39.5)

[a]relevant recorded X-ray in phase one prior to index consultation

Table 5-3: Frequency of recording of holistic assessment and core management processes in phase one.

		Pain assessment ≤14 days *n* (%)	Pain assessment phase 1 *n* (%)	Function assessment ≤14 days *n* (%)	Function assessment phase 1 *n* (%)	Exercise assessment ≤14 days *n* (%)	Exercise assessment phase 1 *n* (%)	Exercise advice ≤14 days *n* (%)	Exercise advice phase 1 *n* (%)	Weight advice ≤14 days *n* (%)	Weight advice phase 1 *n* (%)
Overall		0 (0)	6 (0.2)	0 (0)	11 (0.3)	0 (0)	9 (0.3)	3 (0.1)	3 (0.1)	1 (0.1)	1 (0.1)
Diagnostic group	Joint pain	-	4 (0.2)	-	4 (0.2)	-	6 (0.3)	2 (0.1)	2 (0.1)	0 (0.0)	0 (0.0)
	OA	-	2 (0.2)	-	7 (0.7)	-	3 (0.3)	1 (0.1)	1 (0.1)	1 (0.2)	1 (0.2)
Sex	Female	-	4 (0.2)	-	6 (0.3)	-	3 (0.2)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Male	-	2 (0.2)	-	5 (0.4)	-	6 (0.5)	3 (0.2)	3 (0.2)	1 (0.1)	1 (0.1)
Age	45-64	-	4 (0.3)	-	0 (0.0)	-	2 (0.1)	1 (0.1)	1 (0.1)	1 (0.1)	1 (0.1)
	65-74	-	1 (0.1)	-	1 (0.1)	-	3 (0.4)	1 (0.1)	1 (0.1)	0 (0.0)	0 (0.0)
	75-84	-	1 (0.2)	-	6 (1.0)	-	2 (0.3)	1 (0.2)	1 (0.2)	0 (0.0)	0 (0.0)
	85+	-	0 (0.0)	-	4 (1.8)	-	2 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Site	Knee	-	1 (0.1)	-	6 (0.4)	-	3 (0.2)	3 (0.2)	3 (0.2)	1 (0.1)	1 (0.1)
	Hip	-	1 (0.1)	-	2 (0.3)	-	2 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Ankle/Foot	-	1 (0.5)	-	0 (0.0)	-	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Wrist/Hand	-	1 (0.4)	-	0 (0.0)	-	2 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Unspecified	-	0 (0.0)	-	1 (0.7)	-	1 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Multiple	-	2 (0.6)	-	2 (0.6)	-	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
BMI Category	BMI <25	-	4 (0.9)	-	1 (0.2)	-	3 (0.7)	0 (0.0)	0 (0.0)	-	-
	BMI 25 to <30	-	0 (0.0)	-	3 (0.4)	-	1 (0.1)	1 (0.1)	1 (0.1)	-	-
	BMI 30+	-	2 (0.2)	-	3 (0.3)	-	3 (0.3)	0 (0.0)	0 (0.0)	1 (0.1)	1 (0.1)
	Unknown	-	0 (0.0)	-	4 (0.4)	-	2 (0.2)	2 (0.2)	2 (0.2)	-	-

(con't)

		Pain assessment ≤14 days *n* (%)	Pain assessment phase 1 *n* (%)	Function assessment ≤14 days *n* (%)	Function assessment phase 1 *n* (%)	Exercise assessment ≤14 days *n* (%)	Exercise assessment phase 1 *n* (%)	Exercise advice ≤14 days *n* (%)	Exercise advice phase 1 *n* (%)	Weight advice ≤14 days *n* (%)	Weight advice phase 1 *n* (%)
BNF chapter count	0-4	-	2 (0.1)	-	1 (0.1)	-	4 (0.3)	1 (0.1)	1 (0.1)	0 (0.0)	0 (0.0)
	5-9	-	3 (0.3)	-	3 (0.3)	-	2 (0.2)	2 (0.2)	2 (0.2)	1 (0.2)	1 (0.2)
	10+	-	1 (0.2)	-	7 (1.1)	-	3 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
X-ray, phase 1[a]	No		4 (0.1)		11 (0.4)		7 (0.2)	3 (0.1)	3 (0.1)	1 (0.1)	1 (0.1)
	Yes	-	2 (0.9)	-	0 (0.0)	-	2 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Multiple consultations	Single	-	4 (0.2)	-	2 (0.1)	-	4 (0.2)	3 (0.2)	3 (0.2)	0 (0.0)	0 (0.0)
	Multiple	-	2 (0.2)	-	9 (0.7)	-	5 (0.4)	0 (0.0)	0 (0.0)	1 (0.1)	1 (0.1)
Staff index consultation count	At or below the median	-	0 (0.0)	-	0 (0.0)	-	1 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Above the median	-	6 (0.2)	-	11 (0.4)	-	8 (0.3)	3 (0.1)	3 (0.1)	1 (0.1)	1 (0.1)

[a]relevant recorded X-ray in phase one prior to index consultation

The main analysis examined processes of care within 14 days of a clinical OA consultation; a sensitivity analysis considered processes at any point within phase one.

Due to low levels of recording of these processes within 14 days of a consultation for clinical OA, no logistic regression analysis to assess associations with baseline characteristics was possible for any of these outcomes.

Patient assessment indicators: weight records and X-ray use

10% of patients consulting for clinical OA in phase one had a recorded weight within 14 days of such a consultation (Table 5-4). Achievement of weight recording considered across the whole of phase one of a clinical OA consultation was substantially better at 49.9%. Compared to age band 45-64, patients in age bands 75-84 (adjusted OR 0.64, 95% CI 0.44,0.93) and 85+ (adjusted OR 0.31, 95% CI 0.15,0.63) had significantly lower odds of weight recording and those with an unknown BMI status before the index consultation also had lower adjusted odds at 0.41 (95% CI 0.26,0.64) compared to BMI known not to be overweight. Those with multiple OA or joint pain consultations had greater odds of a weight measurement within 14 days (adjusted OR 2.38, 95% CI 1.83,3.10), compared to single consulters. Estimates of the VPC suggested that 7% of variation was explained by unobserved clinician-level factors in the null model, reducing to 3% in the adjusted model.

18.5% of clinical OA consulters had an X-ray within 14 days of their index consultation. X-ray use across the whole of phase one was slightly greater at 25.8% overall. People with an OA diagnosis had reduced odds of an X-ray within 14 days of a clinical OA consultation compared to joint pain (adjusted OR 0.54, 95% CI 0.41,0.71), although the unadjusted proportion of people with a recorded X-ray was greater in the OA group. Patients in the oldest age band (85+) had reduced odds of a recorded X-ray compared to the youngest (45-64) (adjusted OR 0.48, 95% CI 0.28,0.82). Compared to the knee, people with hip disease had greater odds of a recorded X-ray (adjusted OR 1.92, 95% CI 1.46,2.51) but those with ankle/foot, hand/wrist disease or unspecified site were less likely to be X-rayed. Other associations with recorded X-ray use were BMI status overweight compared to not overweight (adjusted OR 1.61, 95% CI 1.10,2.34), multiple clinical OA consultation compared to single (adjusted OR 4.99, 95% CI 3.95,6.31), and X-ray use before the index consultation (adjusted OR 1.82, 95% CI 1.30, 2.55). Those with higher levels of morbidity had reduced odds of X-ray use

compared to the lowest (adjusted OR 0.63, 95% CI 0.46, 0.87). Estimates of the VPC suggested that 44% of variation was explained by unobserved clinician-level factors in the null model, reducing to 4% in the adjusted model.

Table 5-4: Weight and X-ray records in phase one.

		Weight record within 14 days *n* (%)	Weight record at any point in phase 1 *n* (%)	X-ray record within 14 days *n* (%)	X-ray record at any point in phase 1 *n* (%)
Overall		314 (10.0)	1569 (49.9)	583 (18.5)	811 (25.8)
Diagnostic group	Joint pain	191 (9.2)	1024 (49.4)	375 (18.1)	505 (24.4)
	OA	123 (11.5)	545 (50.7)	208 (19.4)	306 (28.5)
Sex	Female	175 (9.2)	934 (49.1)	361 (19.0)	504 (26.5)
	Male	139 (11.2)	635 (51.2)	222 (17.9)	307 (24.7)
Age	45-64	162 (11.0)	679 (46.0)	288 (19.5)	399 (27.0)
	65-74	92 (10.9)	461 (54.8)	161 (19.1)	216 (25.7)
	75-84	51 (8.5)	331 (55.4)	110 (18.4)	155 (25.9)
	85+	9 (3.9)	98 (43.0)	24 (10.5)	41 (18.0)
Site	Knee	155 (10.2)	744 (49.1)	276 (18.2)	367 (24.2)
	Hip	56 (8.3)	335 (49.4)	174 (25.7)	255 (37.6)
	Ankle/Foot	16 (7.8)	108 (52.9)	18 (8.8)	30 (14.7)
	Wrist/Hand	24 (8.8)	143 (52.6)	30 (11.0)	44 (16.2)
	Unspecified	13 (8.7)	75 (50.0)	9 (6.0)	14 (9.3)
	Multiple	50 (15.4)	164 (50.5)	76 (23.4)	101 (31.1)
BMI Category	BMI <25	45 (10.0)	255 (56.9)	70 (15.6)	102 (22.8)
	BMI 25 to <30	76 (10.2)	455 (61.1)	157 (21.1)	209 (28.1)
	BMI 30+	145 (16.0)	635 (70.0)	186 (20.5)	264 (29.1)
	Unknown	48 (4.6)	224 (21.4)	170 (16.3)	236 (22.6)
BNF chapter count	0-4	136 (8.6)	648 (41.2)	297 (18.9)	403 (25.6)
	5-9	95 (10.5)	503 (55.6)	176 (19.4)	248 (27.4)
	10+	83 (12.5)	418 (62.8)	110 (16.5)	160 (24.0)
X-ray, phase 1[a]	No	280 (9.6)	1452 (49.8)	479 (16.4)	580 (19.9)
	Yes	34 (14.7)	117 (50.6)	104 (45.0)	231 (100.0)
Multiple consultations	Single	126 (6.6)	934 (49.1)	200 (10.5)	331 (17.4)
	Multiple	188 (15.1)	635 (51.2)	383 (30.9)	480 (38.7)
Staff index consultation count	At or below the median	17 (8.0)	120 (56.3)	55 (25.8)	75 (35.2)
	Above the median	297 (10.1)	1449 (49.4)	528 (18.0)	736 (25.1)

[a]relevant recorded X-ray in phase one prior to index consultation
The main analysis examined processes of care within 14 days of a clinical OA consultation; a sensitivity analysis considered processes at any point within phase one.

Table 5-5: Estimates of associations between weight record and relevant X-ray within 14 days of a clinical OA consultation and the independent variables

	Weight record adjusted OR (95% CI)	Relevant X-ray adjusted OR (95% CI)
Diagnosis OA (*reference: joint pain*)	1.00 (0.75,1.34)	↓ **0.54 (0.41,0.71)**
Sex (*reference*: female)	1.26 (0.98,1.62)	0.88 (0.70,1.10)
Age 65-74 (*reference: 45-64*)	0.89 (0.66,1.19)	1.15 (0.88,1.51)
Age 75-84	↓ **0.64 (0.44,0.93)**	1.07 (0.78,1.47)
Age 85+	↓ **0.31 (0.15,0.63)**	↓ **0.48 (0.28,0.82)**
Hip (*reference*: knee)	0.88 (0.63,1.23)	↑ **1.92 (1.46,2.51)**
Ankle/foot	0.85 (0.48,1.49)	↓ **0.54 (0.31,0.95)**
Wrist/hand	1.05 (0.65,1.69)	↓ **0.55 (0.35,0.88)**
Unspecified	0.91 (0.48,1.71)	↓ **0.39 (0.18,0.86)**
Multisite	1.22 (0.83,1.79)	1.01 (0.70,1.45)
BMI 25 to <30 (*reference: BMI <25*)	0.93 (0.62,1.39)	↑ **1.61 (1.10,2.34)**
BMI 30+	1.40 (0.96,2.05)	1.26 (0.87,1.81)
BMI unknown	↓ **0.41 (0.26,0.64)**	1.09 (0.76,1.57)
5-9 BNF chapters (*reference: 0-4*)	1.05 (0.77,1.42)	0.90 (0.69,1.18)
10+ BNF chapters	1.35 (0.96,1.89)	↓ **0.63 (0.46,0.87)**
Multiple consultations in phase 1 (vs. single)	↑ **2.38 (1.83,3.10)**	↑ **4.99 (3.95,6.31)**
X-ray, phase 1[a]	1.39 (0.89,2.16)	↑ **1.82 (1.30,2.55)**
Above the clinician median index consultation count (*reference: at or below the median*)	1.34 (0.74,2.40)	1.06 (0.69,1.65)
Practice 2 (*reference*: Practice 1)	1.52 (0.73,3.16)	0.90 (0.46,1.78)
Practice 3	↓ **0.53 (0.29,0.97)**	↓ **0.12 (0.07,0.21)**
Practice 4	1.59 (0.98,2.60)	↓ **0.01 (0.00,0.05)**
Practice 5	1.29 (0.77,2.17)	↓ **0.01 (0.00,0.05)**
Practice 6	1.88 (0.96,3.67)	↓ **0.01 (0.00,0.08)**
Practice 7	0.93 (0.53,1.64)	1.42 (0.92,2.19)
Practice 8	1.04 (0.51,2.14)	0.83 (0.46,1.50)
VPC[b] (null model)	0.07 (0.02,0.11)	0.44 (0.33,0.52)
VPC[b] (adj. model)	0.03 (0.00,0.07)	0.04 (0.00,0.07)

[a]relevant recorded X-ray in phase one prior to index consultation; [b]variance partition coefficient. Adjusted for all covariates.

Prescribing assessment using routine prescription data

The prescribing processes are shown grouped into those recommended by NICE as first-line relatively safe options, the second-tier options, and those recommended not to be used.

Paracetamol prescription within 14 days of a clinical OA consultation was recorded at 13.8% overall (Table 5-6). Analysis of paracetamol prescription across the whole of phase one showed that there was a much greater proportion of people prescribed paracetamol once the restriction to 14 days of a clinical OA consultation was removed (29.1% overall). A significant association for paracetamol prescription within 14 days of a clinical OA consultation was found for a formal OA diagnosis compared to joint pain code (adjusted OR 1.58, 95% CI 1.23,2.03). Males had lower odds of a paracetamol prescription (adjusted OR 0.76, 95% CI 0.60,0.95) compared to females. Increasing age had a gradient of increasing odds up to OR 3.75 (95% CI 2.53,5.55) for people in age band 85+, compared to 45-64 years. Site effects varied: people with hip disease had slightly increased odds of prescription compared to knee disease (adjusted OR 1.47, 95% CI 1.04,1.79) whilst those with ankle/foot disease had reduced odds (adjusted OR 0.51, 95% CI 0.27,0.96). Increasing morbidity showed a gradient of increasing odds up to adjusted OR 2.04 (95% CI1.53,2.74) for people prescribed from 10+ BNF chapters compared to 0-4. People with multiple clinical OA consultations in phase one also had increased odds of prescription (adjusted OR 1.83 [95% CI 1.46,2.30] compared to single consultation). Estimates of the VPC suggested that 3% of variation was explained by unobserved clinician-level factors in the null model, reducing to 1% in the adjusted model.

Topical NSAID use within 14 days of a clinical OA consultation was recorded in 16.5% of patients with clinical OA. There was no association between odds of a topical NSAID prescription and receipt of an OA diagnosis compared to joint pain (adjusted OR 0.97, 95% CI 0.75,1.24). Similar to paracetamol, there was an apparent gradient of increasing odds of prescription of a topical NSAID with increasing age (adjusted OR 2.71 [95% CI 1.80,4.07] for people in age band 85+ compared to 45-64 years). Males again had reduced odds of prescription compared to females (adjusted OR 0.67, 95% CI 0.53,0.83). Patients with disease at the hip (adjusted OR 0.34, 95% CI 0.25,0.47) and ankle/foot (adjusted OR 0.53, 95% CI 0.32,0.88) had reduced odds of prescription compared to those with knee disease. Increasing morbidity was associated with increased odds of prescription (adjusted OR 1.82 [95% CI 1.37,2.41] for people prescribed from 10+ BNF chapters compared to 0-

4). Multiple clinical OA consultation was associated with increased odds of topical NSAID prescription (adjusted OR 1.59, 95% CI 1.27,1.98). Estimates of the VPC suggested that 11% of variation was explained by unobserved clinician-level factors in the null model, reducing to 10% in the adjusted model.

The proportions of people receiving either paracetamol or topical NSAIDs within 14 days of a clinical OA consultation was 27.1%), and within phase one overall 48.0%. Findings from the multilevel modelling were very similar to those for topical NSAIDs alone although (similar to paracetamol prescription alone) having a OA diagnosis compared to joint pain was associated with increased odds of prescription of either topical NSAID or paracetamol (adjusted OR 1.27, 95% CI 1.03,1.55). Estimates of the VPC suggested that 5% of variation was explained by unobserved clinician-level factors in both the null and adjusted models.

Table 5-6: Frequency of recommended relatively safe pharmacological options in phase one

		Paracetamol ≤14 days *n* (%)	Paracetamol in phase 1 *n* (%)	Topical NSAID ≤14 days *n* (%)	Topical NSAID in phase 1 *n* (%)	Paracetamol or topical NSAID ≤14 days *n* (%)	Paracetamol or topical NSAID in phase 1 *n* (%)
Overall		435 (13.8)	916 (29.1)	520 (16.5)	1038 (33.0)	853 (27.1)	1511 (48.0)
Diagnostic group	Joint pain	216 (10.4)	507 (24.5)	317 (15.3)	629 (30.4)	479 (23.1)	889 (42.9)
	OA	219 (20.4)	409 (38.1)	203 (18.9)	409 (38.1)	374 (34.8)	622 (57.9)
Sex	Female	294 (15.4)	618 (32.5)	357 (18.8)	707 (37.1)	583 (30.6)	1007 (52.9)
	Male	141 (11.4)	298 (24.0)	163 (13.1)	331 (26.7)	270 (21.8)	504 (40.6)
Age	45-64	100 (6.8)	223 (15.1)	159 (10.8)	326 (22.1)	240 (16.2)	456 (30.9)
	65-74	124 (14.7)	264 (31.4)	177 (21.0)	321 (38.1)	267 (31.7)	465 (55.2)
	75-84	147 (24.6)	292 (48.8)	131 (21.9)	280 (46.8)	243 (40.6)	412 (68.9)
	85+	64 (28.1)	137 (60.1)	53 (23.2)	111 (48.7)	103 (45.2)	178 (78.1)
Site	Knee	189 (12.5)	403 (26.6)	282 (18.6)	520 (34.3)	417 (27.5)	707 (46.6)
	Hip	116 (17.1)	231 (34.1)	60 (8.8)	181 (26.7)	159 (23.5)	323 (47.6)
	Ankle/Foot	12 (5.9)	38 (18.6)	23 (11.3)	57 (27.9)	33 (16.2)	79 (38.7)
	Wrist/Hand	18 (6.6)	50 (18.4)	54 (19.9)	91 (33.5)	67 (24.6)	114 (41.9)
	Unspecified	24 (16.0)	51 (34.0)	29 (19.3)	49 (32.7)	50 (33.3)	80 (53.3)
	Multiple	76 (23.4)	143 (44.0)	72 (22.2)	140 (43.1)	127 (39.1)	208 (64.0)

(con't)

		Paracetamol ≤14 days *n* (%)	Paracetamol in phase 1 *n* (%)	Topical NSAID ≤14 days *n* (%)	Topical NSAID in phase 1 *n* (%)	Paracetamol or topical NSAID ≤14 days *n* (%)	Paracetamol or topical NSAID in phase 1 *n* (%)
BMI Category	BMI <25	65 (14.5)	135 (30.1)	68 (15.2)	138 (30.8)	121 (27.0)	209 (46.7)
	BMI 25 to <30	110 (14.8)	226 (30.3)	128 (17.2)	249 (33.4)	212 (28.5)	367 (49.3)
	BMI 30+	138 (15.2)	297 (32.7)	154 (17.0)	336 (37.0)	265 (29.2)	498 (54.9)
	Unknown	122 (11.7)	258 (24.7)	170 (16.3)	315 (30.1)	255 (24.4)	437 (41.8)
BNF chapter count	0-4	122 (7.8)	257 (16.3)	190 (12.1)	339 (21.5)	285 (18.1)	499 (31.7)
	5-9	154 (17.0)	334 (36.9)	169 (18.7)	354 (39.1)	287 (31.7)	520 (57.5)
	10+	159 (23.9)	325 (48.8)	161 (24.2)	345 (51.8)	281 (42.2)	492 (73.9)
X-ray, phase 1[a]	No	393 (13.5)	836 (28.7)	483 (16.6)	960 (32.9)	783 (26.9)	1390 (47.7)
	Yes	42 (18.2)	80 (34.6)	37 (16.0)	78 (33.8)	70 (30.3)	121 (52.4)
Multiple consultations	Single	193 (10.1)	487 (25.6)	265 (13.9)	584 (30.7)	425 (22.3)	832 (43.7)
	Multiple	242 (19.5)	429 (34.6)	255 (20.5)	454 (36.6)	428 (34.5)	679 (54.7)
Staff index consultation count	At or below the median	25 (11.7)	54 (25.4)	37 (17.4)	63 (29.6)	56 (26.3)	92 (43.2)
	Above the median	410 (14.0)	862 (29.4)	483 (16.5)	975 (33.3)	797 (27.2)	1419 (48.4)

[a] relevant recorded X-ray in phase one prior to index consultation

The main analysis examined processes of care within 14 days of a clinical OA consultation; a sensitivity analysis considered processes at any point within phase one.

Table 5-7: Estimates of associations between recommended relatively safe pharmacological options within 14 days of a clinical OA consultation and the independent variables

	Paracetamol adjusted OR (95% CI)	Topical NSAID adjusted OR (95% CI)	Paracetamol or topical NSAID adjusted OR (95% CI)
Diagnosis OA (*reference*: joint pain)	**↑ 1.58 (1.23,2.03)**	0.97 (0.75,1.24)	**↑ 1.27 (1.03,1.55)**
Sex (*reference*: female)	**↓ 0.76 (0.60,0.95)**	**↓ 0.67 (0.53,0.83)**	**↓ 0.65 (0.54,0.78)**
Age 65-74 (*reference*: 45-64)	**↑ 1.93 (1.44,2.60)**	**↑ 2.45 (1.88,3.19)**	**↑ 2.29 (1.84,2.85)**
Age 75-84	**↑ 3.29 (2.43,4.45)**	**↑ 2.32 (1.72,3.13)**	**↑ 3.04 (2.39,3.88)**
Age 85+	**↑ 3.75 (2.53,5.55)**	**↑ 2.71 (1.80,4.07)**	**↑ 3.75 (2.69,5.23)**
Hip (*reference*: knee)	**↑ 1.37 (1.04,1.79)**	**↓ 0.34 (0.25,0.47)**	**↓ 0.69 (0.55,0.87)**
Ankle/foot	**↓ 0.51 (0.27,0.96)**	**↓ 0.53 (0.32,0.88)**	**↓ 0.51 (0.34,0.79)**
Wrist/hand	0.59 (0.35,1.00)	1.27 (0.89,1.82)	0.99 (0.71,1.37)
Unspecified	1.22 (0.74,2.03)	0.92 (0.56,1.50)	1.22 (0.81,1.83)
Multisite	1.19 (0.84,1.66)	1.03 (0.73,1.46)	1.14 (0.85,1.52)
BMI 25 to <30 (reference: BMI <25)	1.07 (0.75,1.52)	1.16 (0.82,1.65)	1.10 (0.83,1.47)
BMI 30+	1.08 (0.76,1.53)	1.04 (0.73,1.47)	1.09 (0.82,1.44)
BMI unknown	0.93 (0.66,1.32)	1.23 (0.87,1.72)	1.04 (0.79,1.38)
5-9 BNF chapters (*reference*: 0-4)	**↑ 1.62 (1.23,2.13)**	**↑ 1.40 (1.08,1.82)**	**↑ 1.53 (1.24,1.89)**
10+ BNF chapters	**↑ 2.04 (1.53,2.74)**	**↑ 1.82 (1.37,2.41)**	**↑ 2.11 (1.67,2.67)**
Multiple consultations in phase 1 (vs. single)	**↑ 1.83 (1.46,2.30)**	**↑ 1.59 (1.27,1.98)**	**↑ 1.69 (1.41,2.03)**
X-ray, phase 1[a] (vs. none)	0.95 (0.64,1.41)	1.00 (0.66,1.53)	1.00 (0.71,1.40)
Above the clinician median index consultation count (*reference*: at or below the median)	1.01 (0.62,1.65)	0.75 (0.44,1.25)	0.85 (0.56,1.28)
Practice 2 (*reference*: Practice 1)	0.76 (0.41,1.43)	**↑ 4.61 (1.87,11.4)**	**↑ 2.27 (1.15,4.51)**
Practice 3	1.20 (0.82,1.75)	1.20 (0.67,2.14)	1.10 (0.72,1.70)
Practice 4	**↓ 0.53 (0.35,0.81)**	1.21 (0.66,2.24)	0.78 (0.49,1.23)
Practice 5	0.90 (0.59,1.37)	**↑ 2.59 (1.39,4.83)**	**↑ 1.76 (1.09,2.82)**
Practice 6	1.10 (0.62,1.97)	0.74 (0.28,1.99)	0.98 (0.49,1.94)
Practice 7	0.80 (0.51,1.26)	0.68 (0.35,1.32)	0.71 (0.44,1.15)
Practice 8	0.55 (0.29,1.04)	**↑ 2.56 (1.19,5.51)**	1.48 (0.81,2.72)
VPC[b] (null model)	0.03 (0.00,0.06)	0.11 (0.06,0.16)	0.05 (0.02,0.08)
VPC[b] (adj. model)	0.01 (0.00,0.03)	0.10 (0.04,0.15)	0.05 (0.02,0.08)

[a]relevant recorded X-ray in phase one prior to index consultation; [b]variance partition coefficient. Adjusted for all covariates.

Adjunct pharmacological management options

Opioids were relatively frequently prescribed at 33.9% of patients (within 14 days of a clinical OA consultation), compared to 13.8% of patients prescribed paracetamol and 16.5% a topical NSAID (Table 5-8). Across phase one as a whole, the proportion of patients prescribed an opioid increased

to 56.0%. An OA diagnosis compared to joint pain was associated with increased odds of opioid prescription (adjusted OR 1.44, 95% CI 1.19,1.75), as was increased age (adjusted OR 1.72 [95% CI 1.24,2.39] for age 85+ compared to age 45-64 years). People with hip disease were more likely to have an opioid prescription (adjusted OR 1.36 [95% CI 1.11,1.68] compared to knee), whereas ankle/foot (adjusted OR 0.43, 95% CI 0.28,0.65) and wrist/hand (adjusted OR 0.56, 95% CI 0.39,0.79) disease had lower odds of prescription. Patients with obesity had increased odds of an opioid prescription (adjusted OR 1.44 [95% CI 1.10,1.89] compared to not overweight), as did those with an increased morbidity burden (adjusted OR 2.66 [95% CI 2.13,3.34] for those prescribed from 10+ BNF chapters compared to 0-4), and those with multiple clinical OA consultations in phase one compared to single (adjusted OR 3.05, 95% CI 2.57,3.63). Estimates of the VPC suggested that 5% of variation was explained by unobserved clinician-level factors in the null model, reducing to 2% (not statistically significantly different to no effect) in the adjusted model.

Oral NSAID prescription was recorded in 18.0% of people within 14 days of a clinical OA consultation and 54.1% across phase one as a whole. Diagnosis was not significantly associated with odds of prescription. Older age was associated with lower odds of an oral NSAID prescription (adjusted OR 0.22 [95% CI 0.12,0.39] for age 85+ c.f. 45-64 years). The highest total morbidity burden was negatively associated with oral NSAID prescription (adjusted OR 0.73, 95% CI 0.54,0.98) compared to the lowest. Having multiple clinical OA consultations compared to a single consultation was associated with such a prescription (adjusted OR 2.61, 95% CI 2.12,3.22). Estimates of the VPC suggested that 11% of variation was explained by unobserved clinician-level factors in the null model, reducing to 6% in the adjusted model.

The NICE-mandated target of co-prescription of a PPI for all patients with OA who are prescribed an oral NSAID was not achieved. Overall 35.6% of patients prescribed an NSAID within 14 days of a clinical OA consultation were also prescribed a PPI. Addition of other forms of gastroprotection (non-PPI) added only modestly (2%) to the proportions of people prescribed an oral NSAID also prescribed gastroprotection. Across phase one overall, 54% of patients prescribed oral NSAIDs also received a PPI prescription. OA was not associated with odds of a PPI prescription. Increasing age

was positively associated with odds of such a prescription (adjusted OR 2.11 [95% CI 1.33,3.35], 2.16 [95% CI 1.19,3.90], 9.66 [95% CI 2.35,39.6] for age bands 65-74, 75-84, 85+ compared to age 45-64). Being in the highest morbidity band (10+ chapters) was also associated with increased odds (adjusted OR 3.11, 95% CI 1.75,5.55, compared to the lowest total morbidity band), as was having multiple clinical OA consultations (adjusted OR 1.73, 95% CI 1.14,2.63). The analysis including sex*age interaction effects was statistically significant for male*age band 75-84 (adjusted OR 5.84 [95% CI 1.77,19.3]) but not for the other interaction terms (which also had point estimates of OR closer to no effect at 1.32 (for male*age band 65-74) and 1.53 (male*age band 85+). Estimates of the VPC suggested that 6% of variation was explained by unobserved clinician-level factors in the null model, reducing to 4% in the adjusted model.

There was a low level of capsaicin prescription at 1.3% overall within 14 days and only 2.3% across all of phase one. Capsaicin prescription was not confined to its licenced indications for hand and knee OA, with recorded prescription in hip, ankle/foot, unspecified and multisite disease. The low level of prescribing of capsaicin made use of a logistic regression analysis unfeasible.

Low levels of orlistat use were shown in this patient group during this period at 1.5% within 14 days of a clinical OA consultation and 3.9% over phase one as a whole. Again, the low frequency of use of orlistat made use of a logistic regression analysis unfeasible.

Table 5-8: Frequency of adjunct recommended prescribing processes in phase one

		Opioid ≤14 days *n* (%)	Opioid in phase 1 *n* (%)	Oral NSAID ≤14 days *n* (%)	Oral NSAID in phase 1 *n* (%)	PPI ≤14 days *n* (%)	PPI in phase 1 *n* (%)	Non-PPI gastroprotection ≤ 14 days *n* (%)	Non-PPI gastroprotection, phase 1 *n* (%)	Capsaicin ≤14 days *n* (%)	Capsaicin in phase 1 *n* (%)	Orlistat ≤14 days *n* (%)	Orlistat in phase 1 *n* (%)
Overall		1065 (33.9)	1760 (56.0)	565 (18.0)	1703 (54.1)	201 (35.6)	920 (54.0)	12 (2.1)	30 (1.8)	42 (1.3)	72 (2.3)	14 (1.5)	35 (3.9)
Diagnostic group	Joint pain	601 (29.0)	1053 (50.8)	346 (16.7)	1075 (51.9)	115 (33.2)	568 (52.8)	5 (1.4)	17 (1.6)	28 (1.4)	44 (2.1)	8 (1.4)	21 (3.7)
	OA	464 (43.2)	707 (65.8)	219 (20.4)	628 (58.5)	86 (39.3)	352 (56.1)	7 (3.2)	13 (2.1)	14 (1.3)	28 (2.6)	6 (1.7)	14 (4.0)
Sex	Female	678 (35.6)	1112 (58.4)	332 (17.4)	1009 (53.0)	128 (38.6)	572 (56.7)	7 (2.1)	15 (1.5)	25 (1.3)	43 (2.3)	11 (2.0)	28 (5.1)
	Male	387 (31.2)	648 (52.2)	233 (18.8)	694 (55.9)	73 (31.3)	348 (50.1)	5 (2.1)	15 (2.2)	17 (1.4)	29 (2.3)	3 (0.8)	7 (2.0)
Age	45-64	405 (27.4)	720 (48.7)	344 (23.3)	734 (49.7)	95 (27.6)	353 (48.1)	5 (1.5)	6 (0.8)	17 (1.2)	30 (2.0)	12 (2.7)	27 (6.2)
	65-74	314 (37.3)	510 (60.6)	134 (15.9)	510 (60.6)	58 (43.3)	286 (56.1)	3 (2.2)	9 (1.8)	7 (0.8)	16 (1.9)	2 (0.8)	7 (2.6)
	75-84	241 (40.3)	368 (61.5)	72 (12.0)	336 (56.2)	36 (50.0)	207 (61.6)	4 (5.6)	10 (3.0)	12 (2.0)	17 (2.8)	0 (0.0)	1 (0.6)
	85+	105 (46.1)	162 (71.1)	15 (6.6)	123 (53.9)	12 (80.0)	74 (60.2)	0 (0.0)	5 (4.1)	6 (2.6)	9 (3.9)	0 (0.0)	0 (0.0)
Site	Knee	497 (32.8)	802 (52.9)	267 (17.6)	783 (51.6)	92 (34.5)	396 (50.6)	5 (1.9)	11 (1.4)	27 (1.8)	41 (2.7)	8 (1.7)	23 (5.0)
	Hip	267 (39.4)	436 (64.3)	116 (17.1)	370 (54.6)	46 (39.7)	200 (54.1)	1 (0.9)	4 (1.1)	2 (0.3)	4 (0.6)	4 (2.2)	7 (3.9)
	Ankle/Foot	31 (15.2)	86 (42.2)	29 (14.2)	120 (58.8)	11 (37.9)	63 (52.5)	0 (0.0)	3 (2.5)	4 (2.0)	5 (2.5)	1 (1.7)	2 (3.4)
	Wrist/Hand	51 (18.8)	117 (43.0)	43 (15.8)	134 (49.3)	12 (27.9)	70 (52.2)	0 (0.0)	2 (1.5)	2 (0.7)	6 (2.2)	1 (1.7)	2 (3.4)
	Unspecified	55 (36.7)	92 (61.3)	28 (18.7)	89 (59.3)	11 (39.3)	51 (57.3)	0 (0.0)	3 (3.4)	1 (0.7)	4 (2.7)	0 (0.0)	0 (0.0)
	Multiple	164 (50.5)	227 (69.8)	82 (25.2)	207 (63.7)	29 (35.4)	140 (67.6)	6 (7.3)	7 (3.4)	6 (1.8)	12 (3.7)	0 (0.0)	1 (0.9)

(con't)

		Opioid ≤14 days *n* (%)	Opioid in phase 1 *n* (%)	Oral NSAID ≤14 days *n* (%)	Oral NSAID in phase 1 *n* (%)	PPI ≤14 days *n* (%)	PPI in phase 1 *n* (%)	Non-PPI gastroprotection ≤ 14 days *n* (%)	Non-PPI gastroprotection, phase 1 *n* (%)	Capsaicin ≤14 days *n* (%)	Capsaicin in phase 1 *n* (%)	Orlistat ≤14 days *n* (%)	Orlistat in phase 1 *n* (%)
BMI Category	BMI <25	137 (30.6)	215 (48.0)	64 (14.3)	209 (46.7)	21 (32.8)	99 (47.4)	1 (1.6)	2 (1.0)	10 (2.2)	13 (2.9)	-	-
	BMI 25 to <30	236 (31.7)	419 (56.2)	133 (17.9)	422 (56.6)	51 (38.3)	241 (57.1)	3 (2.3)	10 (2.4)	7 (0.9)	18 (2.4)	-	-
	BMI 30+	387 (42.7)	610 (67.3)	162 (17.9)	544 (60.0)	60 (37.0)	307 (56.4)	2 (1.2)	7 (1.3)	13 (1.4)	23 (2.5)	14 (1.5)	35 (3.9)
	Unknown	305 (29.2)	516 (49.4)	206 (19.7)	528 (50.5)	69 (33.5)	273 (51.7)	6 (2.9)	11 (2.1)	12 (1.1)	18 (1.7)	-	-
BNF chapter count	0-4	370 (23.5)	639 (40.6)	318 (20.2)	737 (46.8)	88 (27.7)	306 (41.5)	1 (0.3)	7 (0.9)	7 (0.4)	14 (0.9)	1 (0.3)	5 (1.5)
	5-9	350 (38.7)	583 (64.4)	160 (17.7)	541 (59.8)	64 (40.0)	317 (58.6)	4 (2.5)	11 (2.0)	15 (1.7)	21 (2.3)	7 (2.3)	12 (3.9)
	10+	345 (51.8)	538 (80.8)	87 (13.1)	425 (63.8)	49 (56.3)	297 (69.9)	7 (8.0)	12 (2.8)	20 (3.0)	37 (5.6)	6 (2.3)	18 (6.8)
X-ray, phase 1[a]	No	959 (32.9)	1607 (55.1)	496 (17.0)	1565 (53.7)	171 (34.5)	834 (53.3)	12 (2.4)	30 (1.9)	40 (1.4)	70 (2.4)	13 (1.6)	34 (4.1)
	Yes	106 (45.9)	153 (66.2)	69 (29.9)	138 (59.7)	30 (43.5)	86 (62.3)	0 (0.0)	0 (0.0)	2 (0.9)	2 (0.9)	1 (1.3)	1 (1.3)
Multiple consultations	Single	443 (23.3)	939 (49.3)	241 (12.7)	979 (51.4)	69 (28.6)	505 (51.6)	6 (2.5)	19 (1.9)	22 (1.2)	38 (2.0)	5 (1.0)	20 (3.9)
	Multiple	622 (50.1)	821 (66.2)	324 (26.1)	724 (58.3)	132 (40.7)	415 (57.3)	6 (1.9)	11 (1.5)	20 (1.6)	34 (2.7)	9 (2.3)	15 (3.8)
Staff index consultation count	At or below the median	91 (42.7)	126 (59.2)	41 (19.2)	106 (49.8)	15 (36.6)	60 (56.6)	0 (0.0)	0 (0.0)	1 (0.5)	1 (0.5)	1 (1.3)	3 (3.8)
	Above the median	974 (33.2)	1634 (55.7)	524 (17.9)	1597 (54.5)	186 (35.5)	860 (53.9)	12 (2.3)	30 (1.9)	41 (1.4)	71 (2.4)	13 (1.6)	32 (3.9)

[a] relevant recorded X-ray in phase one prior to index consultation

The main analysis examined processes of care within 14 days of a clinical OA consultation; a sensitivity analysis considered processes at any point within phase one.

Table 5-9: Estimates of associations between adjunct recommended pharmacological options within 14 days of a clinical OA consultation and the independent variables

	Opioid adjusted OR (95% CI)	Oral NSAID adjusted OR (95% CI)	PPI adjusted OR (95% CI)
Diagnosis OA (*reference: joint pain*)	↑ **1.44 (1.19,1.75)**	1.03 (0.81,1.31)	1.06 (0.67,1.68)
Sex (*reference: female*)	0.88 (0.74,1.04)	1.10 (0.90,1.35)	0.71 (0.48,1.05)
Age 65-74 (*reference: 45-64*)	↑ **1.38 (1.13,1.70)**	↓ **0.62 (0.49,0.79)**	↑ **2.11 (1.33,3.35)**
Age 75-84	↑ **1.33 (1.05,1.69)**	↓ **0.46 (0.34,0.62)**	↑ **2.16 (1.19,3.90)**
Age 85+	↑ **1.72 (1.24,2.39)**	↓ **0.22 (0.12,0.39)**	↑ **9.66 (2.35,39.6)**
Hip (*reference: knee*)	↑ **1.36 (1.11,1.68)**	1.06 (0.82,1.38)	1.10 (0.66,1.86)
Ankle/foot	↓ **0.43 (0.28,0.65)**	1.01 (0.65,1.58)	1.23 (0.51,2.97)
Wrist/hand	↓ **0.56 (0.39,0.79)**	1.10 (0.75,1.60)	0.87 (0.40,1.92)
Unspecified	1.20 (0.81,1.80)	1.36 (0.85,2.19)	0.97 (0.39,2.41)
Multisite	1.13 (0.86,1.49)	1.24 (0.89,1.72)	0.65 (0.35,1.21)
BMI 25 to <30 (*reference: BMI <25*)	1.00 (0.76,1.33)	1.26 (0.89,1.79)	1.35 (0.67,2.70)
BMI 30+	↑ **1.44 (1.10,1.89)**	1.14 (0.81,1.61)	1.43 (0.71,2.86)
BMI unknown	1.02 (0.78,1.34)	1.33 (0.96,1.85)	1.41 (0.72,2.75)
5-9 BNF chapters (*reference: 0-4*)	↑ **1.62 (1.32,1.98)**	0.99 (0.78,1.26)	1.33 (0.84,2.11)
10+ BNF chapters	↑ **2.66 (2.13,3.34)**	↓ **0.73 (0.54,0.98)**	↑ **3.11 (1.75,5.55)**
Multiple consultations in phase 1 (*vs. single*)	↑ **3.05 (2.57,3.63)**	↑ **2.61 (2.12,3.22)**	↑ **1.73 (1.14,2.63)**
X-ray, phase 1[a] (*vs. none*)	1.04 (0.76,1.42)	1.38 (0.97,1.96)	1.28 (0.68,2.42)
Above the clinician median index consultation count (*reference: at or below the median*)	0.71 (0.50,1.01)	0.90 (0.58,1.41)	1.03 (0.47,2.29)

(con't)

	Opioid adjusted OR (95% CI)	Oral NSAID adjusted OR (95% CI)	PPI adjusted OR (95% CI)
Practice 2 (*reference*: Practice 1)	↑ **2.42 (1.45,4.06)**	0.90 (0.42,1.92)	0.56 (0.18,1.77)
Practice 3	0.97 (0.69,1.37)	1.19 (0.76,1.85)	0.88 (0.47,1.65)
Practice 4	↓ **0.66 (0.46,0.94)**	0.87 (0.54,1.42)	0.77 (0.38,1.58)
Practice 5	0.95 (0.65,1.38)	↓ **0.35 (0.20,0.63)**	2.19 (0.91,5.28)
Practice 6	1.45 (0.87,2.44)	0.91 (0.44,1.89)	0.86 (0.30,2.46)
Practice 7	↑ **1.81 (1.26,2.60)**	↓ **0.29 (0.16,0.52)**	↑ **2.90 (1.11,7.55)**
Practice 8	1.13 (0.69,1.85)	0.59 (0.29,1.17)	0.92 (0.34,2.54)
VPC[b] (null model)	0.05 (0.02,0.07)	0.11 (0.06,0.16)	0.06 (0.01,0.12)
VPC[b] (adj. model)	0.02 (0.00,0.04)	0.06 (0.02,0.09)	0.04 (0.00,0.10)

[a]relevant recorded X-ray in phase one prior to index consultation; [b]variance partition coefficient.
Adjusted for all covariates.

Non-recommended pharmacological management options

Overall, 14.1% of patients with a recorded relative comorbidity contraindication to oral NSAIDs received a prescription for an oral NSAID within 14 days of a clinical OA consultation in phase one (Table 5-10). In the analysis of prescribing across the whole of phase one, there was a large increase in the proportion of people with a recorded relative contraindication receiving an oral NSAID prescription, at 56.0%. The associations for a prescription for oral NSAIDs in the presence of a recorded relative contraindication were similar as for oral NSAIDs as a whole. Patients with older age had reduced odds of an oral NSAID prescription (diminishing adjusted OR of 0.53 [95% CI 0.37,0.76], 0.37 [0.24,0.56], 0.29 [95% CI 0.15,0.58] for age bands 65-74, 75-84, 85+ compared to age 45-64). Patients with multiple consultations had increased odds of prescription (adjusted OR 2.00, 95% CI 1.44,2.76). Estimates of the VPC suggested that 9% of variation was explained by unobserved clinician-level factors in the null model, reducing to 2% in the adjusted model.

There was a low level of etoricoxib 60mg prescription (0.9% of patients receiving an oral NSAID prescription within 14 days were prescribed etoricoxib 60mg), in line with the NICE recommendation that it not be used as a first-line oral NSAID choice. This proportion is similar across the phase one aggregated prescriptions at 0.8%. It was not feasible with such low frequencies to undertake a logistic regression model.

Only a low level of prescribing of rubefacients within 14 days of a clinical OA consultation was identified (in seven people, or 0.2% of patients). Across phase one as a whole, this increased to 0.7%. Again, it was not feasible with such small frequencies to undertake a logistic regression model.

Table 5-10: Frequency of non-recommended prescribing processes in phase one.

		Oral NSAID despite contra-indication ≤14 days *n* (%)	Oral NSAID despite contra-indication phase 1 *n* (%)	Etoricoxib 60mg, ≤14 days *n* (%)	Etoricoxib 60mg, phase 1 *n* (%)	Rubefacient, ≤14 days *n* (%)	Rubefacient, phase 1 *n* (%)	Glucosamine, ≤14 days *n* (%)	Glucosamine, phase 1 *n* (%)
Overall		226 (14.1)	895 (56.0)	5 (0.9)	13 (0.8)	7 (0.2)	21 (0.7)	18 (0.6)	63 (2.0)
Diagnostic group	Joint pain	124 (12.6)	534 (54.2)	4 (1.2)	11 (1.0)	3 (0.1)	10 (0.5)	14 (0.7)	38 (1.8)
	OA	102 (16.6)	361 (58.9)	1 (0.5)	2 (0.3)	4 (0.4)	11 (1.0)	4 (0.4)	25 (2.3)
Sex	Female	133 (13.8)	521 (54.2)	5 (1.5)	9 (0.9)	6 (0.3)	16 (0.8)	10 (0.5)	39 (2.0)
	Male	93 (14.6)	374 (58.7)	0 (0.0)	4 (0.6)	1 (0.1)	5 (0.4)	8 (0.6)	24 (1.9)
Age	45-64	112 (20.6)	282 (51.9)	1 (0.3)	6 (0.8)	3 (0.2)	9 (0.6)	6 (0.4)	20 (1.4)
	65-74	61 (12.6)	280 (57.6)	3 (2.2)	6 (1.2)	1 (0.1)	5 (0.6)	4 (0.5)	17 (2.0)
	75-84	41 (9.7)	245 (57.9)	1 (1.4)	1 (0.3)	2 (0.3)	5 (0.8)	5 (0.8)	23 (3.8)
	85+	12 (8.2)	88 (59.9)	0 (0.0)	0 (0.0)	1 (0.4)	2 (0.9)	3 (1.3)	3 (1.3)
Site	Knee	99 (13.2)	396 (52.9)	2 (0.7)	5 (0.6)	3 (0.2)	13 (0.9)	9 (0.6)	29 (1.9)
	Hip	46 (13.1)	189 (53.7)	2 (1.7)	3 (0.8)	3 (0.4)	5 (0.7)	3 (0.4)	10 (1.5)
	Ankle/Foot	15 (14.0)	66 (61.7)	0 (0.0)	2 (1.7)	0 (0.0)	0 (0.0)	2 (1.0)	2 (1.0)
	Wrist/Hand	17 (12.5)	73 (53.7)	0 (0.0)	1 (0.7)	0 (0.0)	1 (0.4)	2 (0.7)	10 (3.7)
	Unspecified	8 (13.3)	39 (65.0)	0 (0.0)	0 (0.0)	1 (0.7)	1 (0.7)	1 (0.7)	5 (3.3)
	Multiple	41 (20.9)	132 (67.3)	1 (1.2)	2 (1.0)	0 (0.0)	1 (0.3)	1 (0.3)	7 (2.2)
BMI Category	BMI <25	23 (9.9)	111 (47.8)	0 (0.0)	0 (0.0)	1 (0.2)	3 (0.7)	0 (0.0)	8 (1.8)
	BMI 25 to <30	67 (15.3)	259 (59.0)	2 (1.5)	4 (0.9)	2 (0.3)	5 (0.7)	7 (0.9)	21 (2.8)
	BMI 30+	90 (15.4)	346 (59.1)	2 (1.2)	8 (1.5)	4 (0.4)	9 (1.0)	4 (0.4)	19 (2.1)
	Unknown	46 (13.4)	179 (52.2)	1 (0.5)	1 (0.2)	0 (0.0)	4 (0.4)	7 (0.7)	15 (1.4)

(con't)

		Oral NSAID despite contra-indication ≤14 days *n* (%)	Oral NSAID despite contra-indication phase 1 *n* (%)	Etoricoxib 60mg, ≤14 days *n* (%)	Etoricoxib 60mg, phase 1 *n* (%)	Rubefacient, ≤14 days *n* (%)	Rubefacient, phase 1 *n* (%)	Glucosamine, ≤14 days *n* (%)	Glucosamine, phase 1 *n* (%)
BNF chapter count	0-4	75 (14.7)	228 (44.8)	0 (0.0)	0 (0.0)	1 (0.1)	3 (0.2)	0 (0.0)	16 (1.0)
	5-9	82 (14.1)	333 (57.2)	4 (2.5)	9 (1.7)	1 (0.1)	5 (0.6)	9 (1.0)	24 (2.7)
	10+	69 (13.6)	334 (65.7)	1 (1.1)	4 (0.9)	5 (0.8)	13 (2.0)	9 (1.4)	23 (3.5)
X-ray, phase 1[a]	No	192 (13.1)	816 (55.8)	4 (0.8)	12 (0.8)	7 (0.2)	19 (0.7)	18 (0.6)	61 (2.1)
	Yes	34 (25.0)	79 (58.1)	1 (1.4)	1 (0.7)	0 (0.0)	2 (0.9)	0 (0.0)	2 (0.9)
Multiple consultations	Single	94 (10.1)	505 (54.4)	2 (0.8)	8 (0.8)	3 (0.2)	13 (0.7)	11 (0.6)	39 (2.0)
	Multiple	132 (19.7)	390 (58.1)	3 (0.9)	5 (0.7)	4 (0.3)	8 (0.6)	7 (0.6)	24 (1.9)
Staff index consultation count	At or below the median	14 (13.1)	58 (54.2)	0 (0.0)	2 (1.9)	1 (0.5)	2 (0.9)	0 (0.0)	2 (0.9)
	Above the median	212 (14.2)	837 (56.1)	5 (1.0)	11 (0.7)	6 (0.2)	19 (0.6)	18 (0.6)	61 (2.1)

[a]relevant recorded X-ray in phase one prior to index consultation

The main analysis examined processes of care within 14 days of a clinical OA consultation; a sensitivity analysis considered processes at any point within phase one.

Table 5-11: Estimates of associations between non-recommended pharmacological options within 14 days of a clinical OA consultation and the independent variables

	Oral NSAID where relative contra-indication present adjusted OR (95% CI)
Diagnosis OA (*reference*: joint pain)	1.18 (0.82,1.69)
Sex (*reference*: female)	1.14 (0.83,1.55)
Age 65-74 (*reference*: 45-64)	↓ **0.53 (0.37,0.76)**
Age 75-84	↓ **0.37 (0.24,0.56)**
Age 85+	↓ **0.29 (0.15,0.58)**
Hip (*reference*: knee)	1.03 (0.69,1.53)
Ankle/foot	1.29 (0.69,2.41)
Wrist/hand	1.23 (0.68,2.21)
Unspecified	1.24 (0.54,2.86)
Multisite	1.37 (0.86,2.17)
BMI 25 to <30 (reference: BMI <25)	1.49 (0.88,2.55)
BMI 30+	1.28 (0.76,2.16)
BMI unknown	1.18 (0.67,2.08)
5-9 BNF chapters (*reference*: 0-4)	1.09 (0.75,1.59)
10+ BNF chapters	1.10 (0.73,1.64)
Multiple consultations in phase 1 (vs. single)	↑ **2.00 (1.44,2.76)**
X-ray, phase 1[a] (vs. none)	1.60 (0.99,2.60)
Above the clinician median index consultation count (*reference*: at or below the median)	1.06 (0.55,2.05)
Practice 2 (*reference*: Practice 1)	0.67 (0.26,1.71)
Practice 3	↑ **1.78 (1.09,2.89)**
Practice 4	0.83 (0.43,1.58)
Practice 5	↓ **0.44 (0.23,0.87)**
Practice 6	1.27 (0.63,2.55)
Practice 7	↓ **0.24 (0.10,0.55)**
Practice 8	0.74 (0.34,1.61)
VPC[b] (null model)	0.09 (0.02,0.15)
VPC[b] (adj. model)	0.02 (0.00,0.06)

[a]relevant recorded X-ray in phase one prior to index consultation; [b]variance partition coefficient. Adjusted for all covariates.

There was a low level of prescription recorded (18 patients, or 0.6%). Across phase one as a whole, this increased modestly to 2.0%. Due to small numbers of patients recorded as receiving a glucosamine prescription, logistic regression models were not feasible to undertake.

The exercise referral domain was dominated by physiotherapy referrals which accounted for all of the 6.7% of patients (n=212) identified as being referred for exercise or physiotherapy within 14 days of a consultation for clinical OA. When all such referrals within phase one or within 14 days of a clinical OA consultation were considered, 13.2% of patients were identified as referred. As no additional referrals other than to physiotherapy were identified for exercise within 14 days of a clinical OA consultation, the associations in the adjusted multilevel logistic regression model are shown under physiotherapy, below.

Assessment of indicators for assistive devices, in Table 5-12, has been derived from the codes shown in Appendix E.1. Overall 6.7% were recorded with a physiotherapy referral within 14 days of a clinical OA consultation (no additional patients with other assistive device referrals within 14 days), and 13.2% within phase one overall, of which 12.3% was physiotherapy referral. A PQL1 estimation approach to the multilevel logistic regression analysis of recorded physiotherapy referral was used. Patients with a formal OA diagnosis (N05.. Read code) were less likely to be referred to physiotherapy than people with joint pain diagnosis only (adjusted OR 0.51, 95% CI 0.35,0.75). Ankle/foot (adjusted OR 0.24, 95% CI 0.07,0.79) and wrist/hand disease (adjusted OR 0.48, 95% CI 0.24,0.97) were also associated with reduced odds of referral. Multiple consultation was associated with increased odds of physiotherapy referral (adjusted OR 2.75, 95% CI 2.00,3.79) as was a recorded relevant X-ray before the index consultation (adjusted OR 1.78, 95% CI1.14,2.79). Estimates of the VPC suggested that 21% of variation was explained by unobserved clinician-level factors in the null model, reducing to 3% in the adjusted model.

Table 5-12 shows the OT referrals as a proxy for non-ambulatory assistive devices; very low levels of referral (1.1% within 14 days of a clinical OA consultation and 2.4% across phase one as a whole) were found, with no evidence of a substantial variation in frequency by age or disease type. In an adjusted single-level model, patients with a formal OA diagnosis compared to joint pain had reduced odds of referral (adjusted OR 0.24, 95% CI 0.08,0.66), as did those with unknown BMI status compared to not overweight (adjusted OR 0.18, 95% CI 0.04,0.75). Patients with wrist/hand

disease (adjusted OR 12.9, 95% CI 4.01,41.6), unspecified (adjusted OR 31.3, 95% CI 6.02,163) or multisite disease (adjusted OR 4.83, 95% CI 1.41,16.6) compared to knee had increased odds of OT referral, as did those with multiple clinical OA consultations in phase one compared to single (adjusted OR 5.08, 95% CI1.29,19.9) and those with a recorded X-ray prior to the index consultation compared to none (adjusted OR 5.73, 95% CI 2.22,14.8).

Only one (<0.1%) patient was identified as having a referral for weight management support within 14 days of a clinical OA consultation. Across phase one as a whole, this extended to 18 people (0.6%). No logistic regression analysis was feasible.

If all the primary care referrals (to physiotherapy as the main referral destination, supervised exercise programmes, OT, podiatry, or weight management) were considered jointly, 7.8% of people were referred within 14 days of a clinical OA consultation and 14.3% across phase one as a whole. Diagnosis was not significantly associated with primary are referral but patients were more likely to be referred if there was recorded multisite disease compared to knee (adjusted OR 1.62, 95% CI 1.05,2.51), but those with ankle/foot disease compared to the knee had lower odds of referral (adjusted OR 0.29, 95% CI 0.10,0.84). Patients with unknown BMI status compared to not overweight (adjusted OR 0.64, 95% CI 0.41,0.99) had lower odds of referral; those with multiple consultation for clinical OA in phase one compared to single had increased odds of referral (adjusted OR 3.07, 95% CI 2.26,4.15) as did those with a prior X-ray compared to none (adjusted OR 1.95,1.27,3.01). Estimates of the VPC suggested that 28% of variation was explained by unobserved clinician-level factors in the null model, reducing to 4% in the adjusted model.

Table 5-13 shows the percentage of patients referred to orthopædics (18%), rheumatology (0.2%) and the pain clinic (0.1%) within 14 days of a clinical OA consultation. People with OA were more likely to be referred to secondary care (a composite of rheumatology, orthopædics, and pain clinic referrals) than those with joint pain (adjusted OR 1.31, 95% CI 1.04,1.64, Table 5-14).Patients in the oldest age band were less likely to be referred (adjusted OR 0.34 [95% CI 0.20,0.55] for age 85+ compared to 45-64). Patients with ankle/foot problems (adjusted OR 0.36, 95% CI 0.20,0.64) were

less likely to be referred as were those with wrist/hand (adjusted OR 0.48, 95% CI 0.31,0.76) and unspecified site of disease (adjusted OR 0.23, 95% CI 0.11,0.47). Patients with multiple clinical OA consultation in phase one had increased odds of referral (adjusted OR 3.24, 95% CI 2.64,3.98), as did those with prior X-ray use (adjusted OR 1.67, 95% CI 1.19,2.33). Estimates of the VPC suggested that 3% of variation was explained by unobserved clinician-level factors in the null model, reducing to 1% in the adjusted model.

24.5% of clinical OA consulters were recorded as receiving any relevant referral within 14 days of a clinical OA consultation in phase one, increasing to 35.2% for relevant referrals across phase one.

In the adjusted multilevel model, odds of referral diminished for patients with increasing age, judging by the point estimate of effect, but the relationship was only significant for age 85+ compared to 45-64 (adjusted OR 0.37, 95% CI 0.25,0.57). Patients with disease at one of three sites had reduced odds of referral– compared to the knee, ankle/foot (adjusted OR 0.33, 95% CI 0.20,0.56), wrist/hand (adjusted OR 0.64, 95% CI 0.44,0.91), and unspecified (adjusted OR 0.36, 95% CI 0.20,0.62). Patients with multiple consultations for clinical OA in phase one had increased odds of referral (adjusted OR 3.48, 95% CI 2.89,4.19), as did those with recorded X-ray use before the index consultation (adjusted OR 1.88, 95% CI 1.37,2.58).

Table 5-12: Frequencies of primary care referrals in phase one.

		Physiotherapy or exercise ≤14 days *n* (%)	Physiotherapy or exercise in phase 1 *n* (%)	Physiotherapy ≤14 days *n* (%)	Physiotherapy in phase 1 *n* (%)	Occupational therapy ≤14 days *n* (%)	Occupational therapy in phase 1 *n* (%)	Weight management therapy ≤14 days *n* (%)	Weight management therapy in phase 1 *n* (%)	Any primary care referral *n* (%)	Any primary care referral phase *1 n* (%)
Overall		212 (6.7)	416 (13.2)	212 (6.7)	387 (12.3)	34 (1.1)	77 (2.4)	1 (0.0)	18 (0.6)	245 (7.8)	451 (14.3)
Diagnostic group	Joint pain	139 (6.7)	277 (13.4)	139 (6.7)	255 (12.3)	24 (1.2)	47 (2.3)	1 (0.0)	12 (0.6)	163 (7.9)	294 (14.2)
	OA	73 (6.8)	139 (12.9)	73 (6.8)	132 (12.3)	10 (0.9)	30 (2.8)	0 (0.0)	6 (0.6)	82 (7.6)	157 (14.6)
Sex	Female	134 (7.0)	262 (13.8)	134 (7.0)	249 (13.1)	23 (1.2)	54 (2.8)	1 (0.1)	10 (0.5)	157 (8.2)	295 (15.5)
	Male	78 (6.3)	154 (12.4)	78 (6.3)	138 (11.1)	11 (0.9)	23 (1.9)	0 (0.0)	8 (0.6)	88 (7.1)	156 (12.6)
Age	45-64	118 (8.0)	219 (14.8)	118 (8.0)	209 (14.2)	16 (1.1)	38 (2.6)	1 (0.1)	7 (0.5)	134 (9.1)	243 (16.5)
	65-74	50 (5.9)	101 (12.0)	50 (5.9)	90 (10.7)	7 (0.8)	13 (1.5)	0 (0.0)	7 (0.8)	56 (6.7)	102 (12.1)
	75-84	34 (5.7)	74 (12.4)	34 (5.7)	68 (11.4)	6 (1.0)	18 (3.0)	0 (0.0)	4 (0.7)	40 (6.7)	80 (13.4)
	85+	10 (4.4)	22 (9.6)	10 (4.4)	20 (8.8)	5 (2.2)	8 (3.5)	0 (0.0)	0 (0.0)	15 (6.6)	26 (11.4)
Site	Knee	112 (7.4)	202 (13.3)	112 (7.4)	191 (12.6)	9 (0.6)	18 (1.2)	1 (0.1)	10 (0.7)	120 (7.9)	209 (13.8)
	Hip	48 (7.1)	104 (15.3)	48 (7.1)	95 (14.0)	3 (0.4)	11 (1.6)	0 (0.0)	5 (0.7)	51 (7.5)	102 (15.0)
	Ankle/Foot	3 (1.5)	14 (6.9)	3 (1.5)	12 (5.9)	1 (0.5)	2 (1.0)	0 (0.0)	3 (1.5)	4 (2.0)	14 (6.9)
	Wrist/Hand	10 (3.7)	25 (9.2)	10 (3.7)	23 (8.5)	9 (3.3)	12 (4.4)	0 (0.0)	0 (0.0)	19 (7.0)	33 (12.1)
	Unspecified	2 (1.3)	13 (8.7)	2 (1.3)	9 (6.0)	5 (3.3)	12 (8.0)	0 (0.0)	0 (0.0)	7 (4.7)	18 (12.0)
	Multiple	37 (11.4)	58 (17.8)	37 (11.4)	57 (17.5)	7 (2.2)	22 (6.8)	0 (0.0)	0 (0.0)	44 (13.5)	75 (23.1)

(con't)

		Physiotherapy or exercise ≤14 days *n* (%)	Physiotherapy or exercise in phase 1 *n* (%)	Physiotherapy ≤14 days *n* (%)	Physiotherapy in phase 1 *n* (%)	Occupational therapy ≤14 days *n* (%)	Occupational therapy in phase 1 *n* (%)	Weight management therapy ≤14 days *n* (%)	Weight management therapy in phase 1 *n* (%)	Any primary care referral *n* (%)	Any primary care referral phase *1 n* (%)
BMI Category	BMI <25	36 (8.0)	62 (13.8)	36 (8.0)	58 (12.9)	7 (1.6)	13 (2.9)	0 (0.0)	1 (0.2)	43 (9.6)	68 (15.2)
	BMI 25 to <30	45 (6.0)	98 (13.2)	45 (6.0)	90 (12.1)	10 (1.3)	20 (2.7)	0 (0.0)	2 (0.3)	54 (7.2)	102 (13.7)
	BMI 30+	66 (7.3)	143 (15.8)	66 (7.3)	131 (14.4)	13 (1.4)	24 (2.6)	1 (0.1)	13 (1.4)	79 (8.7)	154 (17.0)
	Unknown	65 (6.2)	113 (10.8)	65 (6.2)	108 (10.3)	4 (0.4)	20 (1.9)	0 (0.0)	2 (0.2)	69 (6.6)	127 (12.2)
BNF chapter count	0-4	114 (7.2)	202 (12.8)	114 (7.2)	192 (12.2)	12 (0.8)	32 (2.0)	0 (0.0)	4 (0.3)	126 (8.0)	222 (14.1)
	5-9	57 (6.3)	111 (12.3)	57 (6.3)	103 (11.4)	9 (1.0)	21 (2.3)	1 (0.1)	7 (0.8)	65 (7.2)	122 (13.5)
	10+	41 (6.2)	103 (15.5)	41 (6.2)	92 (13.8)	13 (2.0)	24 (3.6)	0 (0.0)	7 (1.1)	54 (8.1)	107 (16.1)
X-ray, phase 1[a]	No	176 (6.0)	350 (12.0)	176 (6.0)	324 (11.1)	30 (1.0)	69 (2.4)	1 (0.0)	16 (0.5)	205 (7.0)	382 (13.1)
	Yes	36 (15.6)	66 (28.6)	36 (15.6)	63 (27.3)	4 (1.7)	8 (3.5)	0 (0.0)	2 (0.9)	40 (17.3)	69 (29.9)
Multiple consultations	Single	79 (4.1)	181 (9.5)	79 (4.1)	163 (8.6)	9 (0.5)	27 (1.4)	0 (0.0)	8 (0.4)	88 (4.6)	187 (9.8)
	Multiple	133 (10.7)	235 (18.9)	133 (10.7)	224 (18.0)	25 (2.0)	50 (4.0)	1 (0.1)	10 (0.8)	157 (12.7)	264 (21.3)
Staff index consultation count	At or below the median	13 (6.1)	33 (15.5)	13 (6.1)	30 (14.1)	5 (2.3)	9 (4.2)	0 (0.0)	2 (0.9)	18 (8.5)	35 (16.4)
	Above the median	199 (6.8)	383 (13.1)	199 (6.8)	357 (12.2)	29 (1.0)	68 (2.3)	1 (0.0)	16 (0.5)	227 (7.7)	416 (14.2)

[a]relevant recorded X-ray in phase one prior to index consultation

The main analysis examined processes of care within 14 days of a clinical OA consultation; a sensitivity analysis considered processes at any point within phase one.

Table 5-13: Frequencies of secondary care referrals in phase one

		Rheumatology ≤14 days *n* (%)	Rheumatology in phase 1 *n* (%)	Orthopædics ≤14 days *n* (%)	Orthopædics in phase 1 *n* (%)	Pain clinic ≤14 days *n* (%)	Pain clinic in phase 1 *n* (%)	Any secondary care referral *n* (%)	Any secondary care referral phase *1 n* (%)	Any referral ≤14 days *n* (%)	Any referral phase *1 n* (%)
Overall		7 (0.2)	57 (1.8)	565 (18.0)	759 (24.1)	4 (0.1)	22 (0.7)	574 (18.3)	804 (25.6)	772 (24.5)	1107 (35.2)
Diagnostic group	Joint pain	5 (0.2)	38 (1.8)	332 (16.0)	470 (22.7)	1 (0.0)	9 (0.4)	336 (16.2)	497 (24.0)	473 (22.8)	696 (33.6)
	OA	2 (0.2)	19 (1.8)	233 (21.7)	289 (26.9)	3 (0.3)	13 (1.2)	238 (22.2)	307 (28.6)	299 (27.8)	411 (38.3)
Sex	Female	3 (0.2)	36 (1.9)	324 (17.0)	457 (24.0)	2 (0.1)	16 (0.8)	328 (17.2)	489 (25.7)	459 (24.1)	689 (36.2)
	Male	4 (0.3)	21 (1.7)	241 (19.4)	302 (24.3)	2 (0.2)	6 (0.5)	246 (19.8)	315 (25.4)	313 (25.2)	418 (33.7)
Age	45-64	4 (0.3)	33 (2.2)	292 (19.8)	388 (26.3)	1 (0.1)	11 (0.7)	295 (20.0)	414 (28.0)	399 (27.0)	563 (38.1)
	65-74	1 (0.1)	16 (1.9)	166 (19.7)	226 (26.8)	2 (0.2)	4 (0.5)	169 (20.1)	236 (28.0)	214 (25.4)	305 (36.2)
	75-84	2 (0.3)	7 (1.2)	85 (14.2)	113 (18.9)	1 (0.2)	4 (0.7)	88 (14.7)	121 (20.2)	125 (20.9)	187 (31.3)
	85+	0 (0.0)	1 (0.4)	22 (9.6)	32 (14.0)	0 (0.0)	3 (1.3)	22 (9.6)	33 (14.5)	34 (14.9)	52 (22.8)
Site	Knee	2 (0.1)	18 (1.2)	310 (20.4)	398 (26.3)	1 (0.1)	8 (0.5)	312 (20.6)	408 (26.9)	405 (26.7)	546 (36.0)
	Hip	4 (0.6)	17 (2.5)	129 (19.0)	176 (26.0)	2 (0.3)	5 (0.7)	135 (19.9)	190 (28.0)	180 (26.5)	263 (38.8)
	Ankle/Foot	0 (0.0)	2 (1.0)	14 (6.9)	25 (12.3)	0 (0.0)	2 (1.0)	14 (6.9)	29 (14.2)	18 (8.8)	38 (18.6)
	Wrist/Hand	0 (0.0)	5 (1.8)	25 (9.2)	45 (16.5)	0 (0.0)	1 (0.4)	25 (9.2)	50 (18.4)	44 (16.2)	72 (26.5)
	Unspecified	0 (0.0)	7 (4.7)	8 (5.3)	20 (13.3)	1 (0.7)	1 (0.7)	9 (6.0)	28 (18.7)	16 (10.7)	40 (26.7)
	Multiple	1 (0.3)	8 (2.5)	79 (24.3)	95 (29.2)	0 (0.0)	5 (1.5)	79 (24.3)	99 (30.5)	109 (33.5)	148 (45.5)

(con't)

		Rheumatology ≤14 days *n* (%)	Rheumatology in phase 1 *n* (%)	Orthopaedics ≤14 days *n* (%)	Orthopaedics in phase 1 *n* (%)	Pain clinic ≤14 days *n* (%)	Pain clinic in phase 1 *n* (%)	Any secondary care referral *n* (%)	Any secondary care referral phase *1 n* (%)	Any referral ≤14 days *n* (%)	Any referral phase *1 n* (%)
BMI Category	BMI <25	3 (0.7)	8 (1.8)	58 (12.9)	77 (17.2)	1 (0.2)	4 (0.9)	62 (13.8)	85 (19.0)	95 (21.2)	133 (29.7)
	BMI 25 to <30	2 (0.3)	20 (2.7)	143 (19.2)	186 (25.0)	1 (0.1)	4 (0.5)	146 (19.6)	198 (26.6)	190 (25.5)	272 (36.5)
	BMI 30+	1 (0.1)	17 (1.9)	183 (20.2)	252 (27.8)	1 (0.1)	11 (1.2)	184 (20.3)	266 (29.3)	250 (27.6)	358 (39.5)
	Unknown	1 (0.1)	12 (1.1)	181 (17.3)	244 (23.3)	1 (0.1)	3 (0.3)	182 (17.4)	255 (24.4)	237 (22.7)	344 (32.9)
BNF chapter count	0-4	2 (0.1)	27 (1.7)	279 (17.7)	361 (22.9)	1 (0.1)	3 (0.2)	282 (17.9)	377 (24.0)	379 (24.1)	529 (33.6)
	5-9	4 (0.4)	17 (1.9)	174 (19.2)	231 (25.5)	2 (0.2)	6 (0.7)	178 (19.7)	244 (27.0)	229 (25.3)	322 (35.6)
	10+	1 (0.2)	13 (2.0)	112 (16.8)	167 (25.1)	1 (0.2)	13 (2.0)	114 (17.1)	183 (27.5)	164 (24.6)	256 (38.4)
X-ray, phase 1[a]	No	6 (0.2)	53 (1.8)	487 (16.7)	665 (22.8)	4 (0.1)	21 (0.7)	496 (17.0)	710 (24.4)	666 (22.9)	976 (33.5)
	Yes	1 (0.4)	4 (1.7)	78 (33.8)	94 (40.7)	0 (0.0)	1 (0.4)	78 (33.8)	94 (40.7)	106 (45.9)	131 (56.7)
Multiple consultations	Single	0 (0.0)	21 (1.1)	205 (10.8)	340 (17.9)	3 (0.2)	12 (0.6)	208 (10.9)	361 (19.0)	290 (15.2)	504 (26.5)
	Multiple	7 (0.6)	36 (2.9)	360 (29.0)	419 (33.8)	1 (0.1)	10 (0.8)	366 (29.5)	443 (35.7)	482 (38.8)	603 (48.6)
Staff index consultation count	At or below the median	1 (0.5)	5 (2.3)	32 (15.0)	43 (20.2)	1 (0.5)	1 (0.5)	34 (16.0)	47 (22.1)	48 (22.5)	72 (33.8)
	Above the median	6 (0.2)	52 (1.8)	533 (18.2)	716 (24.4)	3 (0.1)	21 (0.7)	540 (18.4)	757 (25.8)	724 (24.7)	1035 (35.3)

[a]relevant recorded X-ray in phase one prior to index consultation

The main analysis examined processes of care within 14 days of a clinical OA consultation; a sensitivity analysis considered processes at any point within phase one.

Table 5-14: Estimates of associations (OR, 95%CI) between referrals within 14 days of a clinical OA consultation and the independent variables.

	Physiotherapy referral adjusted OR[a] (95% CI)	OT referral adjusted OR[b] (95% CI)	Any primary healthcare team referral adjusted OR (95% CI)	Any secondary care referral adjusted OR (95% CI)	Any referral adjusted OR (95% CI)
Diagnosis OA (*reference*: joint pain)	↓ **0.51 (0.35,0.75)**	↓ **0.24 (0.08,0.66)**	0.44 (0.31,0.64)	↑ **1.31 (1.04,1.64)**	0.97 (0.79,1.20)
Sex (*reference*: female)	0.87 (0.64,1.18)	1.12 (0.46,2.71)	0.85 (0.63,1.14)	1.18 (0.97,1.44)	1.07 (0.89,1.28)
Age 65-74 (*reference*: 45-64)	0.85 (0.58,1.23)	0.72 (0.26,2.01)	0.81 (0.56,1.16)	0.92 (0.72,1.16)	0.88 (0.71,1.10)
Age 75-84	0.81 (0.52,1.28)	0.65 (0.20,2.05)	0.81 (0.53,1.24)	0.59 (0.44,0.79)	0.64 (0.49,0.83)
Age 85+	0.52 (0.25,1.07)	1.53 (0.41,5.76)	0.67 (0.36,1.26)	↓ **0.34 (0.20,0.55)**	↓ **0.37 (0.25,0.57)**
Hip (*reference*: knee)	0.99 (0.68,1.44)	0.91 (0.21,3.89)	0.96 (0.67,1.40)	1.01 (0.80,1.29)	1.05 (0.84,1.31)
Ankle/foot	↓ **0.24 (0.07,0.79)**	1.35 (0.15,12.45)	↓ **0.29 (0.10,0.84)**	↓ **0.36 (0.20,0.64)**	↓ **0.33 (0.20,0.56)**
Wrist/hand	↓ **0.48 (0.24,0.97)**	↑ **12.9 (4.01,41.6)**	0.94 (0.55,1.63)	↓ **0.48 (0.31,0.76)**	↓ **0.64 (0.44,0.91)**
Unspecified	0.30 (0.07,1.28)	↑ **31.3 (6.02,163)**	1.12 (0.48,2.62)	↓ **0.23 (0.11,0.47)**	↓ **0.36 (0.20,0.62)**
Multisite	1.40 (0.88,2.21)	↑ **4.83 (1.41,16.6)**	↑ **1.62 (1.05,2.51)**	0.78 (0.57,1.07)	0.87 (0.65,1.17)
BMI 25 to <30 (reference: BMI <25)	0.68 (0.42,1.11)	0.67 (0.20,2.18)	0.70 (0.45,1.11)	1.35 (0.95,1.90)	1.16 (0.86,1.57)
BMI 30+	0.75 (0.47,1.20)	0.66 (0.21,2.11)	0.77 (0.50,1.19)	1.23 (0.88,1.73)	1.14 (0.85,1.54)
BMI unknown	0.70 (0.44,1.11)	↓ **0.18 (0.04,0.75)**	↓ **0.64 (0.41,0.99)**	1.15 (0.83,1.61)	1.00 (0.75,1.34)
5-9 BNF chapters (*reference*: 0-4)	0.82 (0.56,1.19)	1.53 (0.54,4.32)	0.85 (0.59,1.22)	1.04 (0.82,1.32)	0.99 (0.80,1.23)
10+ BNF chapters	0.79 (0.51,1.22)	2.32 (0.79,6.81)	0.94 (0.62,1.41)	0.98 (0.74,1.30)	1.04 (0.81,1.33)
Multiple consultations in phase 1 (vs. single)	↑ **2.75 (2.00,3.79)**	↑ **5.08 (1.29,19.9)**	↑ **3.07 (2.26,4.15)**	↑ **3.24 (2.64,3.98)**	↑ **3.48 (2.89,4.19)**
X-ray, phase 1[c] (vs. none)	↑ **1.78 (1.14,2.79)**	↑ **5.73 (2.22,14.8)**	↑ **1.95 (1.27,3.01)**	↑ **1.67 (1.19,2.33)**	↑ **1.88 (1.37,2.58)**
Above the clinician median index consultation count (*reference*: at or below the median)	1.23 (0.64,2.34)	0.80 (0.02,35.52)	1.03 (0.57,1.86)	1.25 (0.82,1.92)	1.24 (0.85,1.81)

(con't)

	Physiotherapy referral adjusted OR[a] (95% CI)	OT referral adjusted OR[b] (95% CI)	Any primary healthcare team referral adjusted OR (95% CI)	Any secondary care referral adjusted OR (95% CI)	Any referral adjusted OR (95% CI)
Practice 2 (*reference*: Practice 1)	0.75 (0.35,1.61)	Omitted (collinearity)	0.65 (0.30,1.44)	0.93 (0.52,1.65)	0.83 (0.52,1.34)
Practice 3	↓ **0.16 (0.08,0.32)**	Omitted (collinearity)	↓ **0.14 (0.07,0.28)**	0.71 (0.48,1.04)	↓ **0.39 (0.28,0.55)**
Practice 4	↓ **0.02 (0.00,0.11)**	Omitted (collinearity)	↓ **0.05 (0.02,0.15)**	↑ **1.67 (1.18,2.36)**	0.80 (0.59,1.10)
Practice 5	↓ **0.00 (0.00,0.00)**	Omitted (collinearity)	↓ **0.04 (0.01,0.14)**	0.71 (0.47,1.07)	↓ **0.36 (0.25,0.52)**
Practice 6	0.48 (0.21,1.06)	Omitted (collinearity)	0.48 (0.22,1.06)	1.33 (0.78,2.27)	0.91 (0.57,1.45)
Practice 7	↓ **0.40 (0.22,0.71)**	Omitted (collinearity)	↓ **0.41 (0.23,0.72)**	1.11 (0.75,1.65)	0.76 (0.54,1.07)
Practice 8	↓ **0.07 (0.02,0.32)**	Omitted (collinearity)	↓ **0.12 (0.04,0.38)**	0.90 (0.53,1.53)	↓ **0.51 (0.32,0.83)**
VPC[d] (null model)	0.21 (0.12,0.29)	-	0.28 (0.16,0.37)	0.03 (0.00,0.05)	0.05 (0.02,0.07)
VPC[d] (adj. model)	0.03 (0.00,0.07)	-	0.04 (0.00,0.08)	0.01 (0.00,0.03)	0.01 (0.00,0.02)

Adjusted two-level (patients within clinicians) logistic regression model, PQL2 approximation (patients within index clinician) except for [a]adjusted two-level model with PQL1 approximation and [b]adjusted single-level model (practice and clinician accounted for through dummy variables); [c]relevant recorded X-ray in phase one prior to index consultation; [d]variance partition coefficient

5.4 Discussion

5.4.1 Main findings and comparison with previous literature

Of the processes of care for OA in general practice, prescription and referral data seemed to be well-recorded. Conversely, and consistent with the CiPCA analysis described in Chapter Three, section 3.4, coded information about assessment (pain, function) and advice (education, exercise, weight loss) appeared very infrequently in the general practice routinely-recorded EHR. Of the fifteen quality indicators set out in the systematic review in Chapter Two, it is concluded that six indicators at a population-level may be wholly or partially measurable from the routine record, consistent with the findings from the CiPCA analysis in Chapter Three: (i) referral for physiotherapy (though the exercise advice component of this domain seems less measurable from the routine record); (ii) referral for ambulatory assistive devices (if physiotherapy referral is taken as a proxy); (iii) referral for (non-ambulatory) assistive devices (if OT referral is taken as a proxy); (iv) paracetamol prescription; (v) oral NSAID prescription; and (vi) PPI prescription. The supplementary topical NSAID prescription indicator proposed in the review also seems measurable. Other processes of care recommended or referred to in the NICE guidance that could be measured at a population level include capsaicin prescription and orlistat prescription, though these are recorded at such low levels that drawing conclusions about factors associated with their use is not feasible. Those processes recommended not to be undertaken (rubefacient, glucosamine, or etoricoxib prescription) could also be assessed descriptively. X-ray use could also be assessed at a population level though the appropriateness or otherwise of individual investigations could not be determined.

Associations across all measures

The main feature generally associated with increased odds of an indicator being met (in both a desirable sense, such as for the assessment or recommended management indicators, and undesirable, such as for use of oral NSAIDs in the context of a relative contraindication or use of X-rays) is multiple clinical OA consultation in phase one compared to single. The effect of multiple

consultation was discussed in Chapter Four, section 4.4.1and the issue of surveillance bias may also apply to quality achievement, as may clinical severity being linked to both frequency of consultation and more assiduous treatment.[274] It may also reflect a different attitude to medical care in patients consulting more frequently (potentially more 'consumerist'[275]).

For the prescribing measures, an OA diagnosis compared to joint pain was associated with paracetamol prescription and for opioids but not topical or oral NSAIDs, nor PPIs. This contrasted with the lack of any significant association between analgesic type and diagnostic code in the analysis of influences on diagnostic code by Jordan et al. (for which the candidate was a co-author).[17] Male sex was associated with lower odds of prescription of paracetamol and topical NSAIDs. Increasing age was associated with all prescriptions except oral NSAIDs (which showed reducing odds with increasing age). Increasing levels of morbidity measured by the BNF chapter count were associated with increased odds of recorded prescription for all agents except oral NSAIDs, which showed reduced odds in the highest morbidity band. Some of the association between morbidity and pharmacological processes of care may be explained by the fact that a prescription of a drug from the groups defined would increase the number of identified BNF chapters used. Another explanation may be the improvement in quality of care for people with higher levels of multimorbidity as has previously been shown in vulnerable elders by Min et al.[276]

For referral measures, an OA diagnosis compared to joint pain was associated with reduced odds of a primary care referral but increased odds of a secondary care referral. Site effects were mixed, though ankle/foot and wrist/hand disease compared to the knee was associated with reduced odds of physiotherapy or secondary care referral, whilst wrist/hand, unspecified and multisite disease compared to the knee was associated with increased odds of OT referral. A recorded X-ray before the index consultation in phase one compared to none was associated with increased odds of all referrals.

Measure-specific associations

For the recording of weight within 14 days of a clinical OA consultation, the rate of achievement at 10% overall in people known to be overweight was lower than expected given advice about weight reduction as a core component of the NICE OA management guidance.[5] The reduced odds of a weight record with increased age may have been due to the requirement to tackle other issues in a consultation or to greater frailty in the very old making a weight measurement less physically practicable. The reduced odds associated with unknown prior BMI status suggests that some people are less likely to be weighed at any point, possibly as their appearance indicates that they are not overweight, or due to patient preference. It has previously been identified that people are more likely to have a BMI recorded if they are overweight.[253]

The National Collaborating Centre for Chronic Conditions argue for a diagnosis of OA made on clinical grounds without recourse to use of imaging techniques and highlight the "unreliable" nature of radiographic reports as a basis for referral.[4] In this context, the frequency of recorded X-rays within 14 days of a clinical OA consultation (18.5%,) was unexpected. Various studies have highlighted that the grade of OA affects both patient and clinician willingness to consider surgical intervention.[4] The outcome of recorded X-ray use within 14 days of a clinical OA consultation was associated with reduced odds of a recorded OA diagnostic code. This may be due to clinicians using X-rays diagnostically, converse to the NICE guidance, especially given the association between prior X-ray use and an OA diagnostic code as discussed in Chapter Four. The hip was associated with greater odds of X-ray use compared to the knee, whilst the ankle/foot, wrist/hand and unspecified sites were associated with reduced odds. Although no prior analysis of the use of X-rays in diagnosing hip OA was identified in the literature, this may have reflected a perception of hip OA as more difficult to diagnose since the joint cannot so easily be inspected or palpated, or due to difficulty in separating hip and back symptoms. Included in the model was X-ray use recorded before the index consultation. This showed a positive association, which may suggest that some patients receive X-rays relatively frequently. However, due to the assignation of X-rays relative to the index consultation, there was a risk for people consulting multiple times in phase one (those

with a joint pain consultation resulting in an X-ray followed later in phase one by an OA consultation) that some X-rays would have been counted in both the predictor and outcome variables.

Comparison with achievement rates from other studies

Rates of adherence to the assessment (pain, function) and non-pharmacological management indicators (assessment, education and advice) as derived from routinely recorded Read codes was very poor at zero to 0.3%, depending upon the eligibility period used. Compared to rates identified from other studies as shown in Table 5-15 (some of which used self-report or analysis of narrative information within medical records), the conclusion is that recording practice for these indicators could be much improved.

Recorded prescription of paracetamol within 14 days of a clinical OA consultation is lower than other estimates of use (Table 5-15). This may be due to the restriction to 14 days for eligibility in this study. All prevalent cases of clinical OA were included (nearly 40% of phase one consulters had consulted about OA or joint pain in the two years before phase one) and not only incident cases, for whom the analgesic options of paracetamol and topical NSAIDs could most strongly be argued. Another explanation may be over-the-counter analgesic use not captured by the prescribing data download.

Topical NSAIDs prescriptions were also lower than other studies' estimates (Table 5-15). Due to their recommended indication being OA of the hand or knee (as well as for the relief of general musculoskeletal pains),[217] it is unsurprising that hip and ankle/foot were associated with reduced odds of prescription.

Relatively high levels of oral NSAID use (33.9% of patients) were identified, with only approximately one third of people prescribed oral NSAIDs also prescribed gastroprotection (35.6%). This level of oral NSAID prescription is a little lower than previous studies. High levels of NSAID use are not necessarily inappropriate, especially given that the analysis does not discriminate between a one-off prescription and continuous use. Lack of routine prescription of gastroprotection given the NICE

Table 5-15: Comparison of measures from MOSAICS phase one with previous literature

Indicator	Kirk 2003[67]	McGlynn 2003[77]	Wenger 2003[78]	Asch 2004[71]	Chodosh 2004[73]	Higashi 2004[75]	Cadogan 2005[72]	Higashi 2005[136]	Ganz 2006[74]	Porcheret 2007[66] (Knee pain)	Steel 2007[68]	Broadbent 2008[69]	Steel 2008[70]	Zingmond 2009[79]	Li 2011[80]	Østerås 2013[82] (patient self-report)	Grønhaug 2014[277] (patient self-report)	THIS ANALYSIS (phase 1)
Assessment																		
Pain	-	57.3[a]	40	80	-	-	-	-	60.6		23 - 30	27	-	-	-	36	64	**0.0**
Function	-	-	-	-	41	-	-	-	-		42 - 44	43	-	-	6.9[b] - 29.2[c]	24	40	**0.0**
Weight recording	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-		**10.0**
X-ray use	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-		**18.5**
Advice																		
Education	-	-	33 - 36	-	-	-	-	-	68.7	16.4[d]	29 - 31	30	17.7	-	-	-		**0.0**
Education (patient self-report)	-	-	-	-	-	-	-	-	-		-	-	77.8	-	-	19 – 24[e]	36-55[e]	**-**
Exercise	-	57.3[a]	-	27	-	-	-	-	-	46.3	-	-	-	-	25.2	49	84	**0.1**
Weight loss	-	-	-	-	-	-	-	-	-	38.8	-	-	-	-	25	34	40	**0.1**
NSAID risk advice	-	-	4	-	-	-	-	4	38.6 - 50.0		14 - 20	17	-	-	-	-		**0.0**
Non-pharmacological management																		
Physio referral/exercise 'prescription'	-	-	0 - 16	-	-	-	46	16	44	40.3	-	-	24.8	-	-	43	76	**6.7[f]**
Walking aids										54.2							27	**-**
OT referral	-	-	-	-	-	-	-	-	-	4.5	-	-	-	-	-	-		**1.1[f]**
Appliances										5.0							18	**-**
Weight loss programme	-	-	-	-	-	-	-	-	-		-	-	-	-	-	5	8	**<0.1[f]**

(con't)

Pharmacological management																		
Paracetamol first analgesic	45[g]	57.3[a]	43	73	-	79	26	43	58.7	70.6	45 - 51	48	41.1	-	-	46	72	**13.8[h]**
Paracetamol maximum dose before alternative	-	-	33	-	-	-	37	33	-		3 - 6	5	-	-	-	-		**-**
Topical NSAID	-	-	-	-	-	-	-	-	-	41.8	-	-	-	-	-	-		**16.5[h]**
Oral NSAID	45[g]	-	-	-	-	-	-		-	59.2[i]	58 – 60[e]	59[j]	-	-	-	46	56	**18.0[h]**
PPI used with oral NSAIDs	45[g]	-	11	-	-	-	-	11	27.4		-	-	-	26.6	-	-		**35.6[h]**
Opioid/stronger analgesic	-	-	-	-	-	-	-	-	-	59.2	-	-	-	-	-	37	43	**33.9[h]**
Capsaicin	-	-	-	-	-	-	-	-	-	3.0	-	-	-	-	-	-		**1.3[h]**
Glucosamine	-	-	-	-	-	-	-	-	-	40.3	-	-	-	-	-	-		**0.6[h]**
Rubefacient	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-		**0.2[h]**
Referral once nonsurgical treatment failed	45[g]	-	90	-	-	-	-	90	72.7	24.9[k]	90 - 95	90	35.8	-	-	47	41	**18.3[k]**

[a]composite of 3 indicators; [b]nonambulatory; [c]ambulatory; [d]written information; [e]depending on component of education; [f]recorded referral; [g]composite of 4 indicators - individual indicator pass rates 27 – 100; [g]written information; [h]recorded prescription; [i]non-selective NSAIDs; [j]for ibuprofen or COX2 as first used oral NSAID; [k]total proportion advised referral, not confined to treatment failure

guidance to co-prescribe a PPI with oral NSAIDs is a greater concern. Likewise, the use of oral NSAIDs in the presence of a possible relative contraindication (14.1% of those with a potential contraindication had an oral NSAID prescribed within 14 days) is not necessarily inappropriate, given that the identification of the contraindications in this analysis depends upon relative and not absolute contraindications. This marker is of potentially greater use in comparisons between clinicians and practices rather than as a definitive view of the quality of care in a cross-sectional survey.

Few patients in this phase were recorded as receiving weight management advice compared to other estimates (Table 5-15). Low levels of prescription for orlistat in the management of obesity may reflect a wider belief about obesity as a self-inflicted condition, or one that does not respond well to medical intervention.[272]

The low levels of prescription for capsaicin (1.3%) were not in keeping with the patient preferences identified by Fraenkel et al. in 2004,[278] who reported that 40% of patients with knee OA preferred capsaicin to the alternatives of oral NSAIDs, glucosamine, or opioids. The rate of use in this analysis was comparable with the 3% found by Porcheret.[66]

Low levels of prescription for rubefacients and glucosamine were consistent with NICE guidance, and for the latter were substantially lower than the 40.3% identified by Porcheret,[66] suggesting that either there has been a change in prescribing behaviour since that 2007 survey, during which interval the NICE guidance on OA has been produced and updated, or that patients obtain it without prescription.

The apparent rate of referral for physiotherapy or exercise programmes, at 6.7%, was much lower than the comparator rates shown in Table 5-15. The association between reduced odds of physiotherapy referral and a diagnosis of OA compared to joint pain was notable and may suggest that primary care teams, or patients themselves, do not share the confidence in physiotherapy as an effective intervention expressed in the NICE guidelines or relevant systematic reviews.[5,279,280]

Referrals for assistive device assessment or provision was hard to determine and relies on identification of physiotherapy or OT referral as a proxy, suggesting only a low level of assessment in this analysis. Memel et al. have previously suggested that GPs miss functional limitation in patients with OA, with only 31% of moderate to severely disabled patients with OA of the hip or knee correctly identified as such.[281] Overall, recorded referral to OT seems very much lower than would have been suggested by the other studies. Greater information about functional limitation within the routine EHR would have been required to establish the degree of need that was appropriately identified.

The orthopædics referral rate within 14 days of a clinical OA consultation of 18.0% seen here was rather low compared to other estimates but it is not possible to determine from this dataset which patients would be regarded as needing an onward referral. The optimum level for referral has not been determined. Dawson et al. argued that the incidence rate for OA of the hip and knee requiring surgery could not reliably be calculated.[282] Frankel *et al*. estimated prevalence of health care need for hip replacement at 107 per 1000 for men and 173 per 1000 for women (both ages 35 years and over); incidence was estimated at 2.23 per 1000.[283] These figures would suggest that the referral frequency identified in this study is substantially more than the proportion of the population with a health care need for hip replacement. Even if a generous allowance for knee replacement referral is made, and further allowance for a less than perfectly efficient identification of those with a need for joint arthroplasty in primary care, there still appears to be scope for improved primary care based OA management without referral. The reduced odds in the adjusted model for specialist assessment seen for the oldest age group (85+) is of potential concern, raising the possibility of age-related discrimination in access to services, though the method of adjustment for morbidity may be insufficient and have resulted in residual confounding. A previous assessment of provision of joint arthroplasty compared to need[284] identified an ‘n’ shaped curve of the provision:need ratio plotted against age, with the youngest and oldest not having access to arthroplasty at the same rate as the middle age groups. Referral clearly does not equal access to arthroplasty but is a

necessary prerequisite in the UK, where general practice acts as a gatekeeper to secondary care referral.

In general, no evidence to support any hypothesis of a variable effect of age, sex and BMI on prescribing behaviours across clinicians was found.

The amount of variation in process measures explained at the level of the clinician in the null models was between 3% (paracetamol, secondary care referral) and 44% (X-ray use within 14 days of a clinical OA consultation). After the addition of explanatory variables (in the fully adjusted model), the VPC reduced to between 1% (paracetamol, secondary care referral) and 10% (topical NSAIDs), though again not all were statistically significant.

Practice effects are harder to interpret. There may well be an element of practice culture affecting clinical behaviour, such as with some of the prescribing and referral behaviours found to have a significant effect in the two-level multilevel model. A three-level multilevel model might have identified any practice-level effect, though eight clusters is a small number for such an analysis, and no statistically significant degree of variance at the practice level was explained by such a model in an exploratory analysis. There was no apparent consistency between indicators about which practices were outliers.

5.4.2 Strengths and limitations

The use of the entire population of clinical OA consulters in eight practices with data gathered before exposure to any of the MOSAICS study interventions, or to CiPCA-like coding training, has enabled conclusions to be drawn about recorded care in a natural setting.

There were some limitations to this analysis. There was a risk of minor adjustments to routine practice during phase one in the study practices, due to their recruitment to the MOSAICS study during this time, though it remained essentially usual care for UK general practice.

There were many comparisons made within this analysis and one would expect that some of those identified as statistically significant will have occurred by chance, with the acceptable probability of

a type 1 error kept at 5%, as discussed in Chapter Four, section 4.4.3. Some of the effects were reasonably consistent, such as age, gender, morbidity, and multiple clinical OA consultation, though others were harder to explain, such as the effect of disease site on co-prescription of a PPI with NSAIDs, or practice effects.

It is acknowledged that some evidence of achievement may be present in the free text. This has a lower utility for audit and other quality improvement activity (the need for information accessibility as outlined by Baker[100]).

The use of a 14-day period following a consultation for clinical OA to attribute assessment or clinical management actions to the care of OA was arbitrary but in keeping with previous practice at the RIPCHS.[264] It has the advantage of increasing the sensitivity of the analysis to relevant care but at the expense of reducing its specificity. For example, a patient may consult regarding another matter within that timeframe (though with the median number of consultations per patient in a year of 5.4 (2008 figures[265]), two unrelated consultations within 14 days would seem very unlikely to be a frequent occurrence). Drugs that were available on repeat prescription may have been out of synchronisation with consultations, so a patient taking a drug on a longer term basis may not receive the prescription within a 14 day period after a relevant consultation.

The increase in proportions of patients with recorded processes of care across the whole of phase one, when the 14-day restriction to a clinical OA consultation was removed, suggests that the main analysis may have under-estimated actual levels of achievement, though for many indicators achievement rates remained very suboptimal even with the broader time period analysis.

No analysis of defined daily doses (DDD)[285] of the pharmacological processes of care has been undertaken, and so it is possible that there was a discrepancy between quantities of oral NSAIDS prescribed and the gastroprotection prescribed for some patients taking NSAIDs. An analysis that included an assessment of DDD would provide a better insight into the quality of the gastroprotection provided. Ideally, the primary care record would explicitly link processes of care and prescriptions to the relevant morbidities they treated, which would overcome problems of

repeat prescriptions for conditions not appearing on the same day in the record as a morbidity code. Although this facility exists within some clinical systems, such linkage was not available within the dataset obtained.

Co-therapy and co-status as relative contraindications, through assessment of interacting drugs and cardiac, renal and hepatic function test results, would have potentially widened the group of patients considered to have a relative contraindication to oral NSAID use, but it was considered too complex and insufficiently reliable for this purpose.

The use of recorded X-ray within phase one prior to the index consultation meant that there was a variable time at-risk for patients to have such a recorded X-ray. No prior evidence to suggest an appropriate period for X-ray use to be considered was identified in the literature.

Referrals cannot be directly linked to the consultation with OA or joint pain as the referral reason, if Read coded, is not identifiably present in the extracted data. As referrals may be processed outside of the clinical system (such as via the secure web-based Choose and Book[286] electronic referral system), or at a different time point to the consultation (such as by administrative staff), referrals cannot clearly be linked by date to a consultation either. Therefore, this analysis provides a broad overview of referrals to relevant services and specialities rather than being a definitive assessment of referrals for OA and joint pain.

5.4.3 Conclusion

Overall, this chapter has served to demonstrate that certain elements of quality of care for OA can be derived from the routine record, with prescribing and referral measures showing promise for measurement at a population level. Assessment and advice indicators are less feasible to measure from the routine record. In general, the indicators used are more likely to be relevant for longitudinal monitoring of care at a population level rather than as a device to definitively state the current level of quality achieved in the primary care of OA. The next chapter will examine how the level of quality achieved might be derived from the information coded in the electronic recording template discussed in Chapter Three.

Chapter Six: Quality of care for osteoarthritis measured through the template: a cross-sectional study

6.1 Introduction

This chapter introduces the quality assessment derived from the electronic consultation recording template described in Chapter Three. It will report on the recorded assessment and management of OA in primary care during phase two of the MOSAICS trial (the first six months of consultation template use) in the eight MOSAICS practices, i.e. prior to randomisation.

As shown in Chapter Five, the level of routine recording of some important aspects of quality of care (predominantly patient assessment [pain, function] and provision of education and advice) was very low. The template was introduced to improve identification in patient records of elements of care for OA. The aim of the analysis presented in this chapter is to describe the level of recorded quality of care for OA in general practice following the introduction of the consultation recording template. Chapter Seven will explore whether the template was associated with improvements in routinely recorded quality of care. The work reported in these two chapters comprises part of the study reported in a peer-reviewed publication regarding the template and its effect on recorded quality of care, reproduced in Appendix B.[99]

6.2 Aim

The aims of this analysis were to describe the level of recorded quality of care for OA in general practice using quality indicator data captured through a consultation recording template and to compare this to information about similar aspects of care recorded routinely (outside the template). Variation in recorded quality of care associated with patient characteristics, clinician and clinical OA workload, and practice will also be explored.

6.3 Method

For reference, the timeline of the MOSAICS study can be seen in Figure 1-1, page 20. The analysis reported in this chapter examines recorded quality of care in phase two, 0-6 months after template

installation and before randomisation of practices to the model OA consultation intervention or control arms. Data were collected from the eight MOSAICS practices by the informatics team of the comprehensive research network (CRN),[287] from a specification determined by the candidate, just as for the data used in Chapter Five. The data obtained were cleaned by the informatics team and Professor Jordan.

6.3.1 Denominator population

This analysis was conducted on patients consulting at least once for peripheral joint clinical OA during phase two. All patients with the selected Read codes as described in Chapter Three, section 3.5 were eligible except for those consulting about shoulder or elbow OA as previously discussed in Chapter Five.

It should be noted that all patients consulting and recorded with a relevant OA or joint pain code within phase two should have automatically caused the recording template to trigger upon entry of the Read code to the medical record causing a bespoke trigger marker to be written to the record. However as the results show, this did not always happen and there was a small group of patients with a relevant Read code but no record of a template marker. This affected 6.5% of patients overall. Three practices contributed 72% of the episodes in which the template did not trigger but only 39% of the affected patients (see Appendix G.1). The template non-triggering was presumed to result from a software anomaly. The analysis of quality achieved applied only to those patients identified as having triggered the template.

As in Chapter Four (section 4.2, page 104), patients were allocated to an index consultation and index clinician in phase two. Similarly, BMI status was determined as for Chapter Four but the information used was extended to that prior to the index consultation in phase two rather than phase one.

6.3.2 Quality indicators

The recorded measures to be investigated were taken predominantly from the systematic review in Chapter Two.[107] An additional quality measure was derived from the NICE guidelines on OA[5]

(topical NSAID use). Achievement of the indicators was determined as described in Chapter Three, section 3.4.3.

Table 6-1: Quality indicator domains for template-derived recorded management of OA

Quality indicator	Indicator origin	Assessment data source(s)
Patient assessment		
Assessment of pain	Systematic review	Template
Assessment of function	Systematic review	Template
Weight/BMI record	[Prerequisite for advice indicator]	Routinely recorded or template
Prescribing		
Assessment or advice about paracetamol use	Systematic review	Template
Assessment or advice about topical NSAID use	NICE guidance	Template
Education & advice		
Evidence of education or advice for OA	Systematic review	Template
Evidence of exercise advice	Systematic review	Template
Weight loss advice for people with peripheral joint clinical OA and a BMI ≥25kgm^{-2}	Systematic review	Template
Physiotherapy		
Consideration of physiotherapy referral	Systematic review	Template
All template indicators in people referred		
Completion of all 8 template indicators above for people receiving specialist referral	Systematic review	Template

The template indicators were used to describe the quality of care and were compared with their corresponding routinely-recorded indicators (if any) (Table 3-3). The bespoke Read codes used in the template described in Chapter Three, section 3.4.3, were used to determine the level of recorded quality of care achieved through the template system. The template could be completed at any clinical OA consultation during phase two. The routinely-recorded measures were considered to be applicable to a clinical OA consultation if they were recorded within 14 days of any such consultation in phase two, as used in Chapter Five.

6.3.3 Data analysis

The frequency of triggering of the template was determined for all clinical OA consulters in phase two.

Associations were estimated for each of the independent variables used in Chapter Five. Results are presented within the assessment and management themes of assessment, prescribing, education and advice, and physiotherapy, with an assessment of completion of all 8 template indicators in all clinical OA consulters and in people receiving an onward referral.

Descriptive epidemiology

Baseline characteristics of the population of clinical OA consulters were identified and described in relation to the independent variables.

Quality of care was described initially in terms of the percentage of eligible patients achieving the indicators in phase two (six months after template installation). This was repeated for patients in whom at least one template entry had been made, as a sensitivity analysis, on the basis that clinicians may be more likely to make a template entry in those regarded as having significant clinical OA.

Association between quality achievement and independent variables of interest: multilevel logistic regression

The same independent variables as assessed in Chapter Four and Chapter Five were used to identify associations with recorded quality of care. The main analysis was repeated, restricted to patients with at least one completed template entry (other than weight or BMI) in phase two, to assess the associations in the subpopulation in whom clinicians were most engaged with template use. A further exploratory analysis also included recorded X-ray within 14 days after any clinical OA consultation in phase two as an independent variable given that the decision to refer people for an X-ray is known to be associated with management decisions in knee pain even before the X-ray result is known.[12]

Using achievement criteria for the quality indicators as the outcomes, the independent variables of interest were assessed in a two-level binary logistic regression multilevel model of patients nested within the index clinician, using a second order penalised quasi-likelihood (PQL2) approximation, unless otherwise specified. Although tested (and found not to indicate a large or statistically

significant amount of variation in achievement of the quality indicators at the topmost (practice) level), a three-level model was not considered to be feasible to undertake with only 8 practices and so the practice level was adjusted for by use of dummy variables in the model. As assessed in the associations for coding behaviours (Chapter Four), age-sex interaction was investigated for recorded quality of care. The VPC was estimated as described in Chapter Four, section 4.2.4.

Estimates of associations between the independent variables and routinely-recorded information linked to the same domains as assessed via the template (weight recording, paracetamol and topical NSAID prescription, and physiotherapy referral) are shown in Appendix G.2. Except where specified, the estimates were similar to those in phase one.

The data were analysed in SPSS v21[230] for the descriptive epidemiology and, for the multilevel modelling, in Stata 13.1[233] and MLwiN[273] v 2.34 using the runMLwiN[235] command. Adjusted models account for all covariates. Results are presented as ORs and 95% CI, with estimates of the clinician-level variance partition coefficients for the null and fully adjusted models.

6.4 Results

6.4.1 Eligible population

During the six months of phase two, there were 1851 patients identified as consulting with OA or joint pain. Of these, 1730 (93.5%) triggered the recording template, and six further patients were then excluded from the analysis due to records indicating only shoulder or elbow sites of disease. This left 1724 patients, which is the figure taken as the denominator for quality indicator achievement measurement.

Two-thirds of patients triggered the template on only one day, with a median one template trigger, IQR 1-2 trigger days. For patients in whom the template triggered, there were index consultations with 86 clinicians who had a median number of 14 index consultations (unique patients) each (range 1-82).

The characteristics of patients consulting for defined clinical OA in phase two are shown in Table 6-2; they were very similar to the characteristics of consulters in phase one (Chapter Five, Table 5-2, page 147).

Table 6-2: Characteristics of clinical OA consulters, phase two

		***n* (%)**
Total		1724
Diagnostic group, *n* (%)	OA	582 (33.8)
Sex, *n* (%)	Female	1014 (58.8)
Age band, *n* (%)	45-64	817 (47.4)
	65-74	442 (25.6)
	75-84	349 (20.2)
	85+	116 (6.7)
Site, *n* (%)	Knee	855 (49.6)
	Hip	363 (21.1)
	Ankle/foot	125 (7.3)
	Wrist/hand	152 (8.8)
	Unspecified	99 (5.7)
	Multiple	130 (7.5)
BMI status, *n* (%)	Not overweight	312 (18.1)
	Overweight	525 (30.5)
	Obese	540 (31.3)
	Unknown	347 (20.1)
BNF chapter count group, *n* (%)	0-4	485 (28.1)
	5-9	578 (33.5)
	10+	661 (38.3)
Multiple clinical OA consultations, phase 2, *n* (%)		532 (30.9)
Recorded X-ray in phase 2 prior to the index consultation, *n* (%)		86 (5.0)
Patients with a comorbidity as a relative contraindication to oral NSAIDs in 2 years prior to index consultation, phase 2		799 (46.3)

Weight measurement (denominator for weight-related interventions)

The proportions of males and females with a weight record were similar; males were more frequently recorded as being overweight but not obese. Older people were more frequently recorded as being not overweight and substantially less frequently as obese.

Table 6-3: BMI status prior to the phase two index consultation, stratified by sex and age

	BMI <25 *n* (%)	**BMI 25-29.9 *n* (%)**	**BMI 30+ *n* (%)**	**BMI status not recorded *n* (%)**	**Total *n***
Total	312 (18.1)	525 (30.5)	540 (31.3)	347 (20.1)	1724
Sex					
Male	102 (14.2)	242 (34.1)	221 (31.1)	146 (20.6)	710
Female	211 (20.8)	283 (27.9)	319 (31.5)	201 (19.8)	1014
Age					
45-64	136 (16.6)	207 (25.3)	277 (33.9)	197 (24.1)	817
65-74	67 (15.2)	153 (34.6)	164 (37.1)	58 (13.1)	442
75-84	71 (20.3)	125 (35.8)	88 (25.2)	65 (18.6)	349
85+	38 (32.8)	40 (34.5)	11 (9.5)	27 (23.3)	116

6.4.2 Overall template use in phase two

The levels of recorded template use in patients with clinical OA in phase two are shown in Table 6-4, stratified (for the purposes of comparison) by trial arm as used in the cluster trial described in Chapter Eight. Overall recorded use was 61.9% of patients with at least one template entry, and 19.0% in whom all template indicators were achieved. When restricted to only those patients in whom the template successfully triggered, usage was a little better with two-thirds of patients with at least one entry and 20.4% in whom all template indicators were achieved.

6.4.3 Quality assessment: general observations

As shown in Table 6-5, pain (63.3% of patients) and function (62.1%) assessments had relatively good levels of achievement compared to the other recorded measures. There then appeared to be an ordering effect within the template with diminishing frequency of indicator achievement down to 45.3% for weight loss advice and 35.8% for consideration of physiotherapy referral. Provision of written, as opposed to verbal, information, exercise advice and weight loss advice were substantially lower at 0.8 to 2.7%. There appeared to be differences between practices in achievement for all indicators.

None of the template-derived indicators was found to have a significant age-sex interaction except for exercise advice, described below.

Table 6-4: Template use in phase two, patients with clinical OA

		Practice ID										Overall
		Intervention arm					Control arm					
		1	2	7	8	Total	3	4	5	6	Total	Total
All patients	% patients for whom template fired	97.9	98.9	90.6	97.4	96.6	89.8	81.8	93.5	99.0	89.6	93.4
	% patients with at least 1 template entry	70.6	78.7	60.6	89.6	71.7	56.7	40.1	38.8	86.6	50.0	61.9
	% patients with at least 1 template indicator achieved	70.6	78.7	60.0	89.6	71.6	56.7	40.1	38.8	86.6	50.0	61.8
	% patients with all eligible template indicators achieved	30.7	12.8	7.8	24.3	24.2	7.8	9.7	6.1	49.5	12.7	19.0
Restricted to patients in whom the template fired	% patients with at least 1 template entry	72.1	79.6	66.9	92.0	74.2	63.2	49.0	41.5	87.5	55.8	66.2
	% patients with at least 1 template indicator achieved	72.1	79.6	66.3	92.0	74.1	63.2	49.0	41.5	87.5	55.8	66.2
	% patients with all eligible template indicators achieved	31.4	12.9	8.6	25.0	25.1	8.6	11.9	6.6	50.0	14.2	20.4

6.4.4 Quality assessment: patient assessment indicators

Pain assessment

A record of pain assessment was the most frequently achieved template indicator (63.3%, 95% CI 61.0,65.6). The adjusted multilevel model (Table 6-5) showed significantly increased likelihood of pain assessment indicator achievement for older ages (75-84 years compared to age 45-64: adjusted OR 1.46 (95% CI 1.02,2.10). Reduced odds of pain assessment were identified for patients with ankle/foot site of disease (compared to knee): adjusted OR 0.40, 95% CI 0.24,0.66), and for increased morbidity load (BNF chapter count of 10 or more compared to 0-4) (adjusted OR 0.58. 95% CI 0.41,0.84). Patients with multiple clinical OA consultations compared to single (adjusted OR 3.03, 95% CI 2.24,4.11), and those with an index consultation with a clinician who had more than the median number of index consultations (adjusted OR 2.42, 95% CI 1.11,5.27) both had increased odds of pain assessment.

The VPC was estimated at 43% in a constant-only two-level model. This reduced in the fully-adjusted model to an estimated 36% of the residual variation in the model of recording of pain assessments explained by unobserved clinician-level characteristics.

Function assessment

Function assessment was the second most frequently achieved indicator (62.1%, 95% CI 59.8,64.3). The adjusted multilevel model (Table 6-5) showed similar significant associations for achievement of the function assessment indicator as the pain assessment achievement, though the age association for function was not statistically significant. Ankle/foot site of disease compared to knee (adjusted OR 0.35, 95% CI 0.21,0.57), unknown BMI status compared to not overweight (adjusted OR 0.59, 95% CI 0.40,0.88), and increased morbidity load compared to lowest 0-4 chapters (10 or more drugs adjusted OR 0.58, 95% CI 0.41,0.83 were associated with less likelihood of achievement. Multiple consultations compared to single (adjusted OR 3.11, 95% CI 2.31,4.19), and index consultation with a clinician who had more than the median number of index consultations

Table 6-5: Frequency of assessment achievement, phase two, and estimates of association with the independent variables

	Pain assessment		Function assessment		Weight record within 14 days	
	n (%)	OR (95% CI) adjusted 2-level model	*n* (%)	OR (95% CI) adjusted 2-level model	*n* (%)	OR (95% CI) adjusted 2-level model
Overall	1092 (63.3)		1070 (62.1)		424 (24.6)	
Diagnosis						
Joint pain	694 (60.8)	1.00	683 (59.8)	1.00	255 (22.3)	1.00
OA	398 (68.4)	1.26 (0.93,1.72)	387 (66.5)	1.15 (0.85,1.56)	169 (29.0)	0.95 (0.70,1.30)
Sex						
Female	456 (64.2)	1.00	445 (62.7)	1.00	188 (26.5)	1.00
Male	636 (62.7)	1.09 (0.84,1.41)	625 (61.6)	1.01 (0.78,1.30)	236 (23.3)	1.27 (0.98,1.65)
Age						
45-64	506 (61.9)	1.00	497 (60.8)	1.00	201 (24.6)	1.00
65-74	288 (65.2)	1.21 (0.87,1.67)	286 (64.7)	1.32 (0.96,1.82)	121 (27.4)	1.12 (0.82,1.55)
75-84	226 (64.8)	↑ **1.46 (1.02,2.10)**	219 (62.8)	1.40 (0.98,1.99)	83 (23.8)	0.90 (0.62,1.30)
85+	72 (62.1)	1.35 (0.78,2.35)	68 (58.6)	1.16 (0.67,1.98)	19 (16.4)	0.58 (0.31,1.09)
Site						
Knee	564 (66.0)	1.00	554 (64.8)	1.00	216 (25.3)	1.00
Hip	238 (65.6)	0.92 (0.66,1.27)	228 (62.8)	0.84 (0.61,1.16)	82 (22.6)	1.06 (0.75,1.48)
Ankle/Foot	55 (44.0)	↓ **0.40 (0.24,0.66)**	53 (42.4)	↓ **0.35 (0.21,0.57)**	27 (21.6)	0.97 (0.57,1.65)
Wrist/Hand	85 (55.9)	0.83 (0.53,1.29)	87 (57.2)	0.90 (0.58,1.41)	28 (18.4)	0.69 (0.41,1.14)
Unspecified	54 (54.5)	0.67 (0.39,1.15)	49 (49.5)	0.56 (0.33,0.95)	27 (27.3)	1.19 (0.68,2.08)
Multiple	96 (73.8)	0.66 (0.39,1.12)	99 (76.2)	0.91 (0.53,1.55)	44 (33.8)	1.34 (0.82,2.20)
BMI Category (prior to index consultation, phase 2)						
BMI <25	203 (65.1)	1.00	201 (64.4)	1.00	89 (28.5)	1.00
BMI 25 to <30	352 (67.3)	1.31 (0.90,1.91)	348 (66.5)	1.31 (0.91,1.90)	140 (26.7)	0.78 (0.54,1.11)
BMI 30+	351 (65.2)	1.35 (0.93,1.98)	348 (64.7)	1.29 (0.89,1.88)	163 (30.2)	1.09 (0.76,1.56)
Not recorded	186 (53.0)	0.67 (0.45,1.00)	173 (49.3)	↓ **0.59 (0.40,0.88)**	32 (9.2)	↓ **0.26 (0.16,0.42)**
BNF subchapter count						
0-4	302 (62.3)	1.00	300 (61.9)	1.00	106 (21.9)	1.00
5-9	385 (66.6)	0.90 (0.64,1.26)	368 (63.7)	0.72 (0.52,1.01)	168 (29.1)	1.28 (0.90,1.80)
10+	405 (61.3)	↓ **0.58 (0.41,0.84)**	402 (60.8)	↓ **0.58 (0.41,0.83)**	150 (22.7)	0.77 (0.53,1.12)

(con't)

	Pain assessment		Function assessment		Weight record within 14 days	
	n (%)	OR (95% CI) adjusted 2-level model	*n* (%)	OR (95% CI) adjusted 2-level model	*n* (%)	OR (95% CI) adjusted 2-level model
Single/multiple consultations for clinical OA in phase 2						
Single	696 (58.4)	1.00	677 (56.8)	1.00	260 (21.8)	1.00
Multiple	396 (74.4)	↑ **3.03 (2.24,4.11)**	393 (73.9)	↑ **3.11 (2.31,4.19)**	164 (30.8)	↑ **2.13 (1.61,2.82)**
X-ray use prior to index consultation[a]						
No	918 (63.0)	1.00	897 (61.5)	1.00	351 (24.1)	1.00
Yes	174 (65.4)	1.04 (0.57,1.91)	173 (65.0)	1.37 (0.75,2.49)	73 (27.4)	0.88 (0.49,1.56)
Staff member index consultation count dichotomy						
At or below the median	96 (48.7)	1.00	84 (42.6)	1.00	75 (38.1)	1.00
Above the median	996 (65.2)	↑ **2.42 (1.11,5.27)**	986 (64.6)	↑ **3.28 (1.58,6.82)**	749 (49.1)	1.31 (0.69,2.46)
Practice						
Practice 1	428 (70.3)	1.00	418 (68.6)	1.00	189 (31.0)	1.00
Practice 2	70 (75.3)	1.49 (0.16,13.4)	69 (74.2)	1.94 (0.24,15.5)	37 (39.8)	1.94 (0.50,7.53)
Practice 3	127 (57.7)	0.48 (0.17,1.33)	118 (53.6)	0.48 (0.18,1.24)	31 (14.1)	↓ **0.35 (0.15,0.79)**
Practice 4	87 (43.1)	↓ **0.22 (0.06,0.78)**	86 (42.6)	↓ **0.29 (0.09,0.93)**	61 (30.2)	0.95 (0.39,2.33)
Practice 5	90 (39.3)	↓ **0.08 (0.02,0.30)**	94 (41.0)	↓ **0.11 (0.03,0.38)**	28 (12.2)	↓ **0.30 (0.12,0.77)**
Practice 6	84 (87.5)	1.20 (0.20,7.29)	84 (87.5)	1.66 (0.30,9.22)	29 (30.2)	0.93 (0.27,3.22)
Practice 7	104 (63.8)	0.70 (0.21,2.29)	104 (63.8)	0.86 (0.28,2.60)	12 (7.4)	↓ **0.22 (0.08,0.60)**
Practice 8	102 (91.1)	2.31 (0.34,15.6)	97 (86.6)	2.40 (0.41,14.0)	37 (33.0)	1.52 (0.45,5.14)
VPC[b] estimates						
Null model	-	0.43 (0.32,0.52)	-	0.41 (0.29,0.49)	-	0.25 (0.15,0.33)
Adjusted model	-	0.36 (0.25,0.45)	-	0.33 (0.22,0.41)	-	0.18 (0.09,0.26)

[a]relevant recorded X-ray in phase two prior to index consultation; [b]variance partition coefficient. Adjusted ORs for achievement include all covariates.

compared to those with the median number or below (adjusted OR 3.28, 95% CI 1.58,6.82) were associated with achievement.

The level-2 VPC was estimated 41% in a constant-only two-level model. This reduced in the fully-adjusted model to 33%, approximately in keeping with the 37% estimate for pain assessment.

A cross-tabulation of the achievement of the pain and function indicators (Table 6-6) shows considerable, but not complete, overlap (94.3% agreement) in achievement, indicating that not all patients with function assessment indicator achievement also achieved the pain indicator.

Table 6-6: Pain and function assessment indicator achievement cross-tabulation

		Function assessment		
		No	**Yes**	**Total**
Pain assessment	**No**	594	38	632
	Yes	60	1032	1092
	Total	654	1070	1724

Weight recording

The levels of recording of a weight measurement through the template or within 14 days of a consultation for clinical OA in the routine records in this period were rather modest at 24.6% overall. The point estimates for adjusted odds of weight recording in age bands 75-84 and 85+ compared to 45-64 were a little greater than in phase one at 0.90 and 0.58 compared to 0.64 and 0.31 respectively, though the statistical significance was lost in the phase two analysis (with smaller patient numbers). There was increased odds of a weight recording if a patient had multiple clinical OA consultations compared to single (adjusted OR 2.13, 95% CI 1.61,2.82) and reduced odds if BMI status was unknown prior to the index consultation compared to known not-overweight (adjusted OR 0.26, 95% CI 0.26,0.42). These were similar to the findings for phase one reported in Chapter Five, though in phase two the statistical significance of increased age and reduced odds of a weight record was lost (with a modest shift in the point estimate of odds toward no effect of age: for age 78-84 compared to 45-64, OR in phase two was 0.90 (95% CI 0.62,1.30) and for age 85+ the OR was 0.58 (95% CI 0.31,1.09)).

The VPC was estimated at 25%. This reduced in the fully-adjusted model to an estimated 18% of residual variation explained by unobserved clinician-level factors.

6.4.5 Quality assessment: core management indicators

Education and information provision

Overall, 49.2% were recorded as receiving verbal or written information about OA. The adjusted model showed a significant association between achievement and an OA diagnosis compared to joint pain (OR 1.60, 95% CI 1.19,2.14), age 65-74 (compared to 45-64: OR 1.47, 95% CI 1.08,2.01 – and similar but nonsignificant estimates for ages 75-84 and 85+). Compared to the knee, the ankle/foot site (OR 0.47, 95% CI 0.29,0.78) and unspecified site of disease (OR 0.43, 95% CI 0.25,0.74) had reduced odds of achievement, as did having the highest morbidity burden (BNF chapter count 10+ compared to 0-4 chapters, OR 0.64, 95% CI 0.45,0.90). Multiple consultation compared to single (OR 2.87, 95% CI 2.18,3.77), and consultation with a clinician who undertook more than the median number of index consultations compared to below or at the median (OR 3.16, 95% CI 1.54,6.48) showed increased odds of success.

The VPC in the constant-only model was estimated at 40%. This reduced in the fully-adjusted model to an estimated 29% of residual variation explained by unobserved clinician-level factors.

Exercise advice

50.1% of patients were recorded as having achieved the exercise indicator via the recording template. The adjusted model showed a significant association between indicator achievement and a diagnosis of OA compared to joint pain (OR 1.56, 95% CI 1.18,2.08), age band (compared to 45-64 years) 65-74 (OR 1.36, 95% CI 1.01,1.83). Disease sites (compared to the knee) of the ankle/foot (OR 0.46, 95% CI 0.28,0.74), wrist/hand (OR 0.45, 95% CI 0.29,0.69), or unspecified (OR 0.53, 95% CI 0.32,0.89) were associated with reduced odds of success, as was unknown BMI compared to not overweight (OR 0.67, 95% CI 0.46,0.98), and the highest total morbidity category of BNF chapters 10+ compared to 0-4 (OR 0.65, 95% CI 0.47,0.91). Multiple consultation compared to single (OR 2.20, 95% CI 1.69,2.86), and consultation with a clinician who undertook more than the median

number of index consultations compared to below or at the median (OR 2.43, 95% CI 1.28,4.59) both showed increased odds of success.

Analysis of the random slope models showed some evidence that the association of age with exercise advice varied between clinicians. In particular, clinicians with a greater propensity to offer exercise advice had a still greater tendency to offer it in age band 75-84 than clinicians with a lower propensity to offer such advice. There was no evidence that the association of gender and BMI with exercise advice varied between clinicians.

The VPC in the constant-only model was 35%. This reduced in the fully-adjusted model to an estimated 24% of remaining variation caused by unobserved clinician-level factors.

Weight loss advice in people with BMI ≥25kgm^{-2}

This indicator was deemed applicable only to patients known to be overweight. Even correcting for this, only modest levels of advice were recorded, at 45.2%. The adjusted model showed significantly reduced likelihood of achievement with disease at the wrist/hand compared to knee (lower odds - adjusted OR 0.52, 95% CI 0.29,0.94). Multiple consultation compared to single (adjusted OR 2.00, 95% CI 1.42,2.80) was associated with increased odds of success. Other trends, which did not achieve statistical significance in this sample, were for diminishing odds of advice in the older age bands, in all disease sites compared to the knee, and with increased morbidity load.

The constant-only model VPC for weight loss advice was estimated at 34%. This reduced in the fully-adjusted model to an estimated 21% of residual variation in provision of weight advice to those known to be overweight explained by unobserved clinician-level factors.

Table 6-7: Frequency of core management achievement, phase two, and estimates of association with the independent variables

	Education provision *n* (%)	Education provision OR (95% CI)	Exercise advice *n* (%)	Exercise advice OR (95% CI)	Weight advice[a] *n* (%)	Weight advice[a] OR (95% CI)
Overall	849 (49.2)	-	864 (50.1)	-	482 (44.6)	-
Diagnosis						
Joint pain	513 (44.9)	1.00	522 (45.7)	1.00	284 (40.6)	1.00
OA	336 (57.7)	↑ **1.60 (1.19,2.14)**	342 (58.8)	↑ **1.56 (1.18,2.08)**	198 (52.0)	1.29 (0.91,1.85)
Sex						
Female	357 (50.3)	1.00	361 (50.8)	1.00	273 (44.5)	1.00
Male	492 (48.5)	1.01 (0.79,1.29)	503 (49.6)	1.04 (0.82,1.32)	209 (44.8)	0.96 (0.70,1.30)
Age						
45-64	382 (46.8)	1.00	396 (48.5)	1.00	218 (44.5)	1.00
65-74	240 (54.3)	↑ **1.47 (1.08,2.01)**	246 (55.7)	↑ **1.36 (1.01,1.83)**	149 (46.4)	1.21 (0.83,1.76)
75-84	171 (49.0)	1.30 (0.92,1.83)	173 (49.6)	1.06 (0.76,1.48)	94 (42.7)	0.99 (0.63,1.53)
85+	56 (48.3)	1.36 (0.81,2.31)	49 (42.2)	0.74 (0.44,1.24)	21 (42.9)	0.77 (0.36,1.68)
Site						
Knee	450 (52.6)	1.00	462 (54.0)	1.00	274 (48.2)	1.00
Hip	171 (47.1)	0.81 (0.59,1.10)	178 (49.0)	0.79 (0.58,1.06)	87 (43.1)	0.78 (0.52,1.18)
Ankle/Foot	44 (35.2)	↓ **0.47 (0.29,0.78)**	45 (36.0)	↓ **0.46 (0.28,0.74)**	26 (31.3)	0.54 (0.29,1.01)
Wrist/Hand	63 (41.4)	0.79 (0.51,1.22)	51 (33.6)	↓ **0.45 (0.29,0.69)**	30 (33.0)	↓ **0.52 (0.29,0.94)**
Unspecified	36 (36.4)	↓ **0.43 (0.25,0.74)**	41 (41.4)	↓ **0.53 (0.32,0.89)**	21 (36.2)	0.60 (0.31,1.19)
Multiple	85 (65.4)	0.74 (0.46,1.21)	87 (66.9)	0.90 (0.55,1.46)	44 (57.1)	0.82 (0.44,1.51)
BMI Category (prior to index consultation, phase 2)						
BMI <25	152 (48.7)	1.00	159 (51.0)	1.00	-	-
BMI 25 to <30	284 (54.3)	1.34 (0.94,1.90)	288 (55.1)	1.31 (0.93,1.85)	225 (44.0)	1.00
BMI 30+	281 (52.2)	1.39 (0.97,1.97)	284 (52.8)	1.29 (0.91,1.82)	245 (46.1)	1.29 (0.94,1.76)
Not recorded	132 (37.6)	0.82 (0.56,1.21)	133 (37.9)	↓ **0.67 (0.46,0.98)**	12 (33.3)	-
BNF subchapter count						
0-4	224 (46.2)	1.00	234 (48.2)	1.00	106 (45.5)	1.00
5-9	302 (52.2)	0.88 (0.64,1.20)	307 (53.1)	0.88 (0.64,1.20)	172 (44.6)	0.79 (0.52,1.22)
10+	323 (48.9)	↓ **0.64 (0.45,0.90)**	323 (48.9)	↓ **0.65 (0.47,0.91)**	204 (44.3)	0.71 (0.45,1.11)

(con't)

	Education provision *n* (%)	Education provision OR (95% CI)	Exercise advice *n* (%)	Exercise advice OR (95% CI)	Weight advice[a] *n* (%)	Weight advice[a] OR (95% CI)
Single/multiple consultations for clinical OA in phase 2						
Single	530 (44.5)	1.00	551 (46.2)	1.00	306 (41.2)	1.00
Multiple	319 (60.0)	↑ **2.87 (2.18,3.77)**	313 (58.8)	↑ **2.20 (1.69,2.86)**	176 (52.1)	↑ **2.00 (1.42,2.80)**
X-ray use prior to index consultation[b]						
No	706 (48.4)	1.00	719 (49.3)	1.00	447 (43.9)	1.00
Yes	143 (53.8)	1.06 (0.61,1.85)	145 (54.5)	1.15 (0.66,2.00)	35 (57.4)	1.32 (0.66,2.66)
Staff member index consultation count dichotomy						
At or below the median	55 (27.9)	1.00	59 (29.9)	1.00	33 (27.0)	1.00
Above the median	794 (52.0)	↑ **2.87 (1.28,6.45)**	805 (52.7)	↑ **3.16 (1.54,6.48)**	449 (46.9)	↑ **2.43 (1.28,4.59)**
Practice ID						
Practice 1	366 (60.1)	1.00	359 (58.9)	1.00	104 (54.7)	1.00
Practice 2	58 (62.4)	1.40 (0.22,8.75)	57 (61.3)	1.40 (0.28,7.05)	15 (62.5)	2.55 (0.53,12.4)
Practice 3	55 (25.0)	↓ **0.14 (0.05,0.37)**	63 (28.6)	↓ **0.22 (0.09,0.50)**	13 (22.0)	↓ **0.24 (0.10,0.60)**
Practice 4	74 (36.6)	0.36 (0.12,1.09)	75 (37.1)	0.40 (0.15,1.07)	21 (32.3)	0.45 (0.17,1.19)
Practice 5	65 (28.4)	↓ **0.10 (0.03,0.33)**	68 (29.7)	↓ **0.14 (0.05,0.40)**	8 (12.9)	↓ **0.10 (0.03,0.30)**
Practice 6	80 (83.3)	2.14 (0.45,10.2)	79 (82.3)	2.17 (0.53,8.87)	36 (78.3)	3.02 (0.76,12.0)
Practice 7	66 (40.5)	0.39 (0.14,1.08)	70 (42.9)	0.49 (0.20,1.23)	7 (18.4)	↓ **0.28 (0.11,0.75)**
Practice 8	85 (75.9)	2.07 (0.42,10.2)	93 (83.0)	2.61 (0.62,10.9)	21 (77.8)	2.95 (0.73,11.9)
VPC[c] estimates						
Null model	-	0.40 (0.29,0.49)	-	0.35 (0.24,0.43)	-	0.34 (0.22,0.43)
Adjusted model	-	0.29 (0.18,0.37)	-	0.24 (0.15,0.31)	-	0.21 (0.11,0.30)

[a]denominator 1065 patients known to have BMI ≥25; [b]relevant recorded X-ray in phase two prior to index consultation; [c]variance partition coefficient. Adjusted ORs for achievement include all covariates.

6.4.6 Quality assessment: pharmacological management indicators

Assessment or consideration of paracetamol use

Analysis of responses to the paracetamol assessment indicator demonstrated an overall achievement rate of 56.3%. The adjusted model showed increased odds of achievement of assessment or consideration of paracetamol use (compared to 45-64) in patients in age groups 65-74 (adjusted OR 1.63, 95% CI 1.20,2.22), 75-84 (adjusted OR 1.64, 95% CI 1.15,2.27), and age 85+ (adjusted OR 1.70, 95% CI 1.02,2.83). Patients with ankle/foot disease compared to the knee (adjusted OR 0.44, 95% CI 0.27,0.71), and unknown BMI status compared to not overweight (adjusted OR 0.61, 95% CI 0.42,0.89) had lower odds of achievement. Patients with multiple consultations compared to single (adjusted OR 2.98, 95% CI 2.26,3.94) and consultation with clinicians conducting more than the median number of index consultations compared to at or below the median (adjusted OR 3.18, 95% CI 1.66,6.08) had increased odds of achievement.

As a comparison, paracetamol prescription was recorded in 17.3% of patients within 14 days of a clinical OA consultation during phase two (identified from the routine record).

The VPC was estimated at 35% in the constant-only model. This reduced in the fully-adjusted model to 26% of residual variation estimated to be due to clinician-level factors.

Table 6-8: Frequency of pharmacological management, phase two, and estimates of association with the independent variables

	Paracetamol considered		Topical NSAID considered,	
	n (%)	OR (95% CI)	*n* (%)	OR (95% CI)
Overall	970 (56.3)	-	832 (48.3)	-
Diagnosis				
Joint pain	606 (53.1)	1.00	522 (45.7)	1.00
OA	364 (62.5)	1.32 (0.99,1.76)	310 (53.3)	↑ **1.36 (1.01,1.83)**
Sex				
Female	394 (55.5)	1.00	345 (48.6)	1.00
Male	576 (56.8)	0.90 (0.71,1.14)	487 (48.0)	0.87 (0.68,1.12)
Age				
45-64	425 (52.0)	1.00	380 (46.5)	1.00
65-74	271 (61.3)	↑ **1.63 (1.20,2.22)**	230 (52.0)	↑ **1.38 (1.01,1.88)**
75-84	208 (59.6)	↑ **1.62 (1.15,2.27)**	172 (49.3)	↑ **1.50 (1.06,2.13)**
85+	66 (56.9)	↑ **1.70 (1.02,2.83)**	50 (43.1)	1.22 (0.71,2.10)
Site				
Knee	495 (57.9)	1.00	438 (51.2)	1.00
Hip	218 (60.1)	1.03 (0.76,1.39)	156 (43.0)	↓ **0.68 (0.50,0.93)**
Ankle/Foot	50 (40.0)	↓ **0.44 (0.27,0.71)**	47 (37.6)	↓ **0.56 (0.34,0.93)**
Wrist/Hand	70 (46.1)	0.72 (0.47,1.10)	68 (44.7)	0.86 (0.55,1.34)
Unspecified	47 (47.5)	0.61 (0.37,1.03)	36 (36.4)	↓ **0.43 (0.25,0.75)**
Multiple	90 (69.2)	0.73 (0.44,1.19)	87 (66.9)	0.93 (0.56,1.54)
BNF subchapter count				
BMI <25	187 (59.9)	1.00	139 (44.6)	1.00
BMI 25 to <30	315 (60.2)	1.09 (0.77,1.54)	279 (53.3)	↑ **1.48 (1.04,2.12)**
BMI 30+	314 (58.4)	1.08 (0.76,1.54)	286 (53.2)	↑ **1.71 (1.19,2.45)**
Not recorded	154 (43.9)	↓ **0.61 (0.42,0.89)**	128 (36.5)	0.83 (0.56,1.23)
BMI Category (prior to index consultation, phase 2)				
0-4	248 (51.1)	1.00	216 (44.5)	1.00
5-9	352 (60.9)	1.08 (0.79,1.48)	306 (52.9)	1.01 (0.73,1.39)
10+	370 (56.0)	0.72 (0.51,1.01)	310 (46.9)	↓ **0.69 (0.48,0.97)**
Single/multiple consultations for clinical OA in phase 2				
Single	602 (50.5)	1.00	521 (43.7)	1.00
Multiple	368 (69.2)	↑ **2.98 (2.26,3.94)**	311 (58.5)	↑ **2.75 (2.08,3.65)**
X-ray use prior to index consultation[a]				
No	809 (55.5)	1.00	704 (48.3)	1.00
Yes	161 (60.5)	1.33 (0.76,2.35)	128 (48.1)	1.03 (0.58,1.82)
Staff member index consultation count dichotomy				
At or below the median	70 (35.5)	1.00	58 (29.4)	1.00
Above the median	900 (58.9)	↑ **3.18 (1.66,6.08)**	774 (50.7)	↑ **2.87 (1.28,6.45)**
Practice ID				
Practice 1	387 (63.5)	1.00	352 (57.8)	1.00
Practice 2	51 (54.8)	0.60 (0.12,3.02)	50 (53.8)	0.68 (0.08,5.54)
Practice 3	101 (45.9)	0.48 (0.21,1.13)	61 (27.7)	↓ **0.22 (0.07,0.65)**
Practice 4	83 (41.1)	0.36 (0.13,0.99)	79 (39.1)	0.41 (0.11,1.47)
Practice 5	80 (34.9)	↓ **0.16 (0.06,0.46)**	78 (34.1)	↓ **0.15 (0.04,0.57)**
Practice 6	84 (87.5)	2.87 (0.63,13.1)	77 (80.2)	1.53 (0.25,9.27)
Practice 7	86 (52.8)	0.69 (0.27,1.79)	78 (47.9)	0.48 (0.14,1.60)
Practice 8	98 (87.5)	3.40 (0.72,16.0)	57 (50.9)	0.81 (0.12,5.45)
VPC[b] **estimates**				
Null model	-	0.35 (0.24,0.43)	-	0.41 (0.30,0.49)
Adjusted model		0.26 (0.16,0.33)		0.37 (0.26,0.46)

[a]relevant recorded X-ray in phase two prior to index consultation; [b]variance partition coefficient.
Adjusted ORs for achievement include all covariates.

Assessment or consideration of topical NSAID use

Achievement of the topical NSAID indicator was a little lower than for paracetamol at 48.3% overall. There was considerable overlap (87% agreement) between the assessment and consideration of paracetamol indicator with the topical NSAID indicator as shown in Table 6-9.

Table 6-9: Paracetamol and topical NSAID consideration cross-tabulation

		Achievement of topical NSAID consideration indicator		
		Not achieved	**Achieved**	**Total**
Achievement of paracetamol consideration indicator	**Not achieved**	712	42	754
	Achieved	180	790	970
	Total	892	832	1724

The adjusted model shows increased odds of topical NSAID indicator achievement and OA diagnosis compared to joint pain (OR 1.36, 95% CI 1.01,2.14). Age (compared to 45-64 years) showed increased achievement for patients in age bands 65-74 years (OR 1.38, 95% CI 1.01,1.88) and 75-84 years (OR 1.50, 95% CI 1.06,2.13). Patients with BMI status, compared to not overweight, of overweight (OR 1.48, 95% CI 1.04,2.12) or obese (OR 1.71, 95% CI 1.19,2.45) also had increased odds of achievement, as did those with multiple clinical OA consultations compared to single (OR 2.75, 95% CI 2.08,3.65), and those who had an index consultation with a clinician who undertook more than the median number of index consultations compared to below or at the median (OR 2.87, 95% CI 1.28,6.45). Patients with recorded disease sites, compared to the knee, of the hip (OR 0.68, 95% CI 0.50,0.93), ankle/foot (OR 0.56, 95% CI 0.34,0.93) and unspecified (OR 0.43, 95% CI 0.25,0.75) had reduced odds of achievement, as did those with high levels of total morbidity (BNF chapter count 10+ compared to 0-4 OR 0.69, 95% CI 0.48,0.97).

Topical NSAID prescription was identified from the routine record in 24.9% of people within 14 days of a clinical OA consultation in phase two, compared to 48.3% of patients in whom the topical NSAID indicator was achieved (i.e. topical NSAID use was assessed or considered) through the template.

In the constant-only model, VPC was estimated at 41%. This reduced in the fully-adjusted model to an estimated 37% of residual variance explained by unobserved clinician-level factors.

6.4.7 Quality assessment: physiotherapy consideration and recorded referral

Consideration of physiotherapy

The rate of achievement for the physiotherapy referral indicator was the lowest of the eight indicators at 35.8%. The adjusted model showed significantly lower odds of achievement in patients with unspecified site of disease compared to the knee (OR 0.50, 95% CI 0.28,0.88). Patients with multiple clinical OA consultations compared to single (OR 2.20, 95% CI 1.71,2.84), and consultation with a clinician who undertook more than the median number of index consultations compared to below or at the median (OR 2.49, 95% CI 1.32,4.73) had increased odds of achievement.

For comparison, physiotherapy referral within 14 days of a clinical OA consultation in phase two was recorded within the routine records in 7.2% (95% CI 6.0,8.6) of patients. Estimates of the OR indicated no statistically significant association between receipt of an OA diagnosis and odds of physiotherapy referral in phase two, unlike in phase one.

The constant-only model VPC estimate was 27%. This reduced in the fully-adjusted model to an estimated 20% of residual variation in consideration of physiotherapy referral explained by unobserved clinician-level factors.

6.4.8 Achievement of all template indicators

Achievement of all 8 template-derived indicators

To assess complete achievement of all indicators, patients known not to be overweight or with unknown weight status were measured out of seven indicators, and those known to be overweight or obese out of eight – including weight loss advice. Weight recording within 14 days of a clinical OA consultation was not included. Overall, the achievement rate for all applicable indicators was 20.4% (95% CI 18.5,22.3). The adjusted model (Table 6-11) showed a positive association between achievement of all indicators and an OA diagnosis compared to joint pain (OR 1.84, 95% CI 1.30,2.60), having multiple clinical OA consultation compared to single (OR 2.14, 95% CI 1.56,2.92), and consultation with a clinician who undertook more than the median number of index consultations compared to below or at the median (OR 2.37, 95% CI 1.04,5.42).

Table 6-10: Frequency of physiotherapy indicator achievement, phase two, and estimates of association with the independent variables

	Physiotherapy referral considered	
	n **(%)**	**OR (95% CI)**
Overall	617 (35.8)	-
Diagnosis		
Joint pain	374 (32.7)	1.00
OA	243 (41.8)	↑ **1.34 (1.01,1.78)**
Sex		
Female	263 (37.0)	1.00
Male	354 (34.9)	1.12 (0.89,1.42)
Age		
45-64	296 (36.2)	1.00
65-74	148 (33.5)	0.80 (0.59,1.07)
75-84	129 (37.0)	1.06 (0.76,1.48)
85+	44 (37.9)	1.02 (0.62,1.69)
Site		
Knee	317 (37.1)	1.00
Hip	126 (34.7)	0.97 (0.72,1.30)
Ankle/Foot	33 (26.4)	0.79 (0.48,1.31)
Wrist/Hand	50 (32.9)	1.10 (0.71,1.69)
Unspecified	23 (23.2)	↓ **0.50 (0.28,0.88)**
Multiple	68 (52.3)	1.05 (0.67,1.65)
BMI Category (prior to index consultation, phase 2)		
BMI <25	118 (37.8)	1.00
BMI 25 to <30	200 (38.2)	0.93 (0.66,1.31)
BMI 30+	194 (36.1)	1.03 (0.73,1.46)
Not recorded	105 (29.9)	0.76 (0.52,1.12)
BNF subchapter count		
0-4	167 (34.4)	1.00
5-9	211 (36.5)	0.96 (0.70,1.31)
10+	239 (36.2)	0.86 (0.62,1.21)
Single/multiple consultations for clinical OA in phase 2		
Single	387 (32.5)	1.00
Multiple	230 (43.2)	↑ **2.20 (1.71,2.84)**
X-ray use prior to index consultation[a]		
No	500 (34.3)	1.00
Yes	117 (44.0)	1.26 (0.76,2.12)
Staff member index consultation count dichotomy		
At or below the median	39 (19.8)	1.00
Above the median	578 (37.9)	↑ **2.49 (1.32,4.73)**
Practice ID		
Practice 1	287 (47.1)	1.00
Practice 2	43 (46.2)	1.20 (0.27,5.32)
Practice 3	50 (22.7)	↓ **0.35 (0.16,0.76)**
Practice 4	55 (27.2)	0.42 (0.17,1.06)
Practice 5	33 (14.4)	↓ **0.10 (0.03,0.29)**
Practice 6	54 (56.3)	0.88 (0.24,3.28)
Practice 7	53 (32.5)	0.56 (0.23,1.34)
Practice 8	42 (37.5)	0.53 (0.14,2.02)
VPC[b] **estimate**		
Null model	-	0.27 (0.17,0.34)
Adjusted model	-	0.20 (0.12,0.28)

[a]relevant recorded X-ray in phase two prior to index consultation; [b]variance partition coefficient.
Adjusted ORs for achievement include all covariates.

For the 8 indicators combined, in a null (constant-only) two-level model, VPC was estimated at 33%. This reduced in the fully-adjusted model to an estimated 23% of the residual variation in achievement of all 8 indicators explained by unobserved clinician-level factors.

Achievement of all template-derived indicators prior to referral

There were 358 (20.8%) patients who were referred on for further care within 14 days of a qualifying consultation. Overall, only a modest proportion (26.0%, 95% CI 21.7,30.8) were recorded as having achieved all of the applicable indicators within phase two prior to referral. Although there were small numbers of patients involved, the adjusted model shows a significant association between achievement of all indicators in people who were referred on for further care and multiple clinical OA consultation compared to single (adjusted OR 4.45, 95% CI 2.04,9.70 (Table 6-11).

VPC in a constant-only two-level model of achievement of all 8 template indicators in people referred was estimated at 48%. This reduced in the fully-adjusted model to 36%, somewhat larger than for achievement of the 8 indicators in the whole study population.

6.4.9 Restriction of template derived indicators to people with at least one template entry: a sensitivity analysis

When the number of people with recorded achievement of the template quality indicators was assessed as a proportion of those with at least one template entry recorded in phase two (n=1142), rates of pain and function assessment achievement increased to 95.6% and 93.7% respectively compared to 63.3% and 62.1% respectively for all consulters (Table 6-12). The indicator achievement rate was 54.0% for consideration of physiotherapy referral and 30.7% for completion of all applicable template indicators.

Table 6-11: Frequency of all indicator achievement, phase two, and estimates of association with the independent variables

	All template indicators[a]		Achievement of all template indicators[a] before referral	
	n (%)	OR (95% CI)	*n* (%)	OR (95% CI)
Overall	351 (20.4)		93 (26.0)	
Diagnosis				
Joint pain	189 (16.5)	1.00	57 (25.1)	1.00
OA	162 (27.8)	↑ **1.84 (1.30,2.60)**	36 (27.5)	1.47 (0.63,3.43)
Sex				
Female	155 (21.8)	1.00	49 (28.7)	1.00
Male	196 (19.3)	1.13 (0.84,1.51)	44 (23.5)	0.77 (0.37,1.59)
Age				
45-64	157 (19.2)	1.00	52 (26.7)	1.00
65-74	93 (21.0)	0.94 (0.65,1.35)	23 (25.8)	0.81 (0.32,2.05)
75-84	78 (22.3)	1.15 (0.76,1.74)	14 (23.0)	0.80 (0.26,2.47)
85+	23 (19.8)	0.78 (0.41,1.47)	4 (30.8)	1.16 (0.14,9.71)
Site				
Knee	177 (20.7)	1.00	59 (28.2)	1.00
Hip	65 (17.9)	0.94 (0.65,1.38)	17 (23.0)	0.94 (0.37,2.40)
Ankle/Foot	21 (16.8)	0.96 (0.51,1.79)	2 (20.0)	1.15 (0.13,10.1)
Wrist/Hand	24 (15.8)	0.92 (0.52,1.61)	2 (9.1)	0.39 (0.05,2.88)
Unspecified	16 (16.2)	0.69 (0.35,1.37)	1 (9.1)	1.13 (0.07,17.2)
Multiple	48 (36.9)	1.11 (0.67,1.83)	12 (37.5)	1.53 (0.46,5.10)
BMI Category (prior to index consultation, phase 2)				
BMI <25	69 (22.1)	1.00	11 (20.0)	1.00
BMI 25 to <30	120 (22.9)	0.81 (0.53,1.22)	38 (32.8)	1.09 (0.35,3.39)
BMI 30+	107 (19.9)	0.83 (0.54,1.27)	30 (26.8)	0.87 (0.27,2.77)
Not recorded	55 (15.7)	0.80 (0.50,1.29)	14 (18.7)	0.70 (0.21,2.40)
BNF subchapter count				
0-4	88 (18.1)	1.00	26 (22.4)	1.00
5-9	119 (20.6)	0.95 (0.64,1.40)	33 (27.7)	0.88 (0.35,2.20)
10+	144 (21.8)	0.95 (0.62,1.44)	34 (27.6)	1.19 (0.39,3.68)
Single/multiple consultations for clinical OA in phase 2				
Single	215 (18.0)	1.00	31 (17.4)	1.00
Multiple	136 (25.6)	↑ **2.14 (1.56,2.92)**	62 (34.4)	↑ **4.45 (2.04,9.70)**
X-ray use prior to index consultation[a]				
No	286 (19.6)	1.00	72 (25.5)	1.00
Yes	65 (24.4)	1.28 (0.72,2.29)	21 (27.6)	2.49 (0.77,8.05)
Staff member index consultation count dichotomy				
At or below the median	18 (9.1)	1.00	4 (9.5)	1.00
Above the median	333 (21.8)	↑ **2.37 (1.04,5.42)**	89 (28.2)	5.02 (0.88,28.5)
Practice ID				
Practice 1	191 (31.4)	1.00	48 (32.9)	1.00
Practice 2	12 (12.9)	0.32 (0.06,1.71)	1 (8.3)	0.07 (0.00,4.33)
Practice 3	19 (8.6)	↓ **0.20 (0.07,0.56)**	6 (14.0)	0.30 (0.05,1.85)
Practice 4	24 (11.9)	0.34 (0.12,1.01)	1 (2.9)	0.04 (0.00,1.09)
Practice 5	15 (6.6)	↓ **0.11 (0.03,0.39)**	2 (10.0)	0.08 (0.00,1.49)
Practice 6	48 (50.0)	1.87 (0.45,7.77)	24 (77.4)	7.77 (0.85,71.1)
Practice 7	14 (8.6)	↓ **0.25 (0.08,0.79)**	5 (8.9)	0.18 (0.03,1.34)
Practice 8	28 (25.0)	0.65 (0.15,2.93)	6 (40.0)	1.51 (0.13,17.0)
VPC[b] estimates				
Null model	-	0.33 (0.20,0.42)	-	0.48 (0.24,0.61)
Adjusted model	-	0.23 (0.12,0.32)	-	0.36 (0.10,0.50)

[a]relevant recorded X-ray in phase two prior to index consultation; [b]variance partition coefficient.
Adjusted ORs for achievement include all covariates.

Table 6-12: Quality indicator summary of template-derived achievement, all patients triggering the template and those with at least one entry recorded

	All patients in whom template triggered, phase 2			Patients with at least one template entry, phase 2		
Indicator	**Numerator**	**Denominator (patients eligible for the indicator)**	**% achieving the indicator (95% CI)**	**Numerator**	**Denominator (patients eligible for the indicator)**	**% achieving the indicator (95% CI)**
Pain assessment	1092	1724	63.3 (61.0,65.6)	1092	1142	95.6 (94.3,96.7)
Function assessment	1070	1724	62.1 (59.8,64.3)	1070	1142	93.7 (92.1,95.0)
Education or information provision	849	1724	49.2 (46.9,51.6)	849	1142	74.3 (71.7,76.8)
Exercise advice or provision	864	1724	50.1 (47.8,52.5)	864	1142	75.7 (73.1,78.1)
Weight loss advice for people with BMI ≥25kgm^{-2}	477	1065	44.8 (41.8,47.8)	477	741	65.0 (61.5,68.4)
Assessment of/recommendation for paracetamol therapy	970	1724	56.3 (53.9,58.6)	970	1142	84.9 (82.7,86.9)
Assessment of/recommendation for topical NSAID therapy	832	1724	48.3 (45.9,50.6)	832	1142	72.9 (70.2,75.4)
Consideration of physiotherapy referral	617	1724	35.8 (33.6,38.1)	617	1142	54.0 (51.1,56.9)
Achievement of all applicable indicators[a]	351	1724	20.4 (18.5,22.3)	351	1142	30.7 (28.1,33.5)
Achievement of all applicable indicators in patients recorded as referred within 14 days of a clinical OA consultation, phase 2[a]	93	358	26.0 (21.7,30.8)	93	261	35.6 (30.1,41.6)

[a]Achievement of all applicable indicators refers to all 7 excluding weight in those not known to be overweight or obese, and all 8 indicators in those known to be overweight or obese.

6.4.10 Effects of a decision to request an X-ray following a clinical OA consultation

When an X-ray recorded within 14 days after a clinical OA consultation in phase two was also included as an independent variable in the multilevel models, there was a significant association between such a recorded X-ray and achievement of pain, function, consideration of paracetamol or topical NSAIDs, information provision and exercise advice, with estimated adjusted ORs between 1.75 (95% CI 1.22,2.51) for exercise advice and 2.38 95% CI 1.53,3.68) for pain assessment for a recorded X-ray compared to none).

Table 6-13: Template-derived quality indicator measures, phase two: association of achievement with recorded X-ray within 14 days of a clinical OA consultation

Indicator	X-ray recorded within 14 days of a clinical OA consultation vs. none OR (95% CI)[a]
Pain	↑ **2.38 (1.53,3.68)**
Function	↑ **2.19 (1.44,3.32)**
Weight record	0.86 (0.58,1.27)
Paracetamol considered	↑ **2.35 (1.60,3.45)**
Top. NSAID considered	↑ **2.06 (1.41,3.02)**
Information provision	↑ **1.98 (1.36,2.86)**
Exercise advice	↑ **1.75 (1.22,2.51)**
Weight loss advice	1.34 (0.66,2.68)
Physio considered	1.26 (0.75,2.12)
All eight QI	1.18 (0.65,2.12)
All eight QI in patients with a recorded referral	2.59 (0.80,8.35)

[a]Fully adjusted two-level logistic regression model including recorded X-ray within 14 days after a clinical OA consultation in phase two as well as the other independent variables (not shown)

6.5 Discussion

6.5.1 Main findings and comparison with previous literature

The levels of quality achieved were best for the first two items on the template: pain and function assessment recording, at 63% and 62% achievement respectively. Most of the remaining template-derived indicators were achieved in the 45% to 56% range, with the exception of consideration of physiotherapy referral at 36%. The rates of achievement for the assessment and advice indicators were substantially better than the rates identified in Chapter Five, section 5.3.2 for the baseline recording through the routine EHR (0% to 0.3%) and compare much more favourably with recorded achievement in other studies as shown in Table 5-15, page 181.

The fact that the first two items on the template were the most frequently-completed may have been due to an ordering effect. The relatively low frequency of achievement for the physiotherapy consideration indicator compared to the other template indicators may have been due to its position at the end of the template, but may also have reflected clinician beliefs about the utility of physiotherapy for OA such as those underlying the variation and under-use of exercise in knee pain and OA as described by Cottrell et al.[288] It may also have reflected the availability of physiotherapy services in primary care, as clinicians may not have wished to raise the issue of physiotherapy referral with patients if there was not adequate service provision to meet demand.[289]

Common associations with indicator achievement

For the template-derived quality measures, the most frequent associations with indicator achievement were a diagnosis of OA rather than joint pain, older age rather than age 45-64, the patient making multiple rather than single consultations for clinical OA within phase two, and an index consultation with a clinician who held more than the median number of index consultations, compared to the median or fewer. The most consistent associations with reduced odds of indicator achievement were a disease site of the ankle/foot (and, with less consistency, unspecified site) compared to the knee, unknown BMI status compared to not overweight, a higher total morbidity measure based on a BNF subchapter count of 10 or more compared to 0-4 chapters, and lower achievement by two practices (3 and 5) compared to practice 1.

Comparison of template-derived and routinely-recorded achievements

When comparing recorded achievement of consideration and use of paracetamol or topical NSAID prescription between the template-derived assessment and recorded prescription, it was apparent that the quality achievement data for assessment from the template show substantially higher achievement than that the recorded actual prescription in the routine record. Paracetamol consideration was achieved in 56.3% of clinical OA consulters in phase two, whereas 17.3% were recorded as receiving a prescription within 14 days of a clinical OA consultation. Similarly, for topical NSAIDs, 48.3% of patients were recorded as having been assessed for these, whilst 24.9% were recorded as receiving a prescription within 14 days of a clinical OA consultation in phase two. It may

be that the relatively safe options had previously been tried in some patients and found to be ineffective. Estimation of ORs for the effects of the independent variables for the prescribing indicators measured in the template and routine record showed similar trends for age, though the effect of increasing age on actual prescription was substantially greater than on the consideration indicator. Odds of recorded prescription were increased by the morbidity burden as measured by BNF chapter count.

Physiotherapy consideration via the template was substantially more frequent at 35.8% compared to the recorded actual referral within 14 days of a clinical OA consultation at 7.2%. The independent variables showed a positive association for patients with an OA diagnosis compared to joint pain in the template physiotherapy consideration, unlike for the recorded referral (which showed lower odds of referral for people with OA in phase one and no difference in phase two), and likewise for the staff index consultation count above the median compared to at or below the median. The wrist/hand site compared to the knee was associated with reduced odds of actual referral, whereas unspecified site was associated with reduced odds of recorded consideration in the template. Actual referral to physiotherapy, rather than recorded consideration, may depend on factors such as health beliefs and access, as noted above. Further work to understand the origin (patient or clinician) and cause of the discrepancy would be required. It was notable that there was no association between consideration of physiotherapy use or referral and BNF chapter count, unlike for many pharmacological options, suggesting that clinicians or patients tend to increase pharmacological options rather than make use of other strategies.

An assessment of random effects of age, sex, and BMI across clinicians was made, on the basis that studies have previously identified age and sex discrimination,[269-271] and as previously noted, some clinicians are prone to negative stereotypes about obesity.[272] However, no strong or consistent evidence was found to support any hypothesis of a variable effect across clinicians of age, sex or BMI on recorded quality of care behaviours.

Increased odds of template indicator achievement or of a record of recommended management options for people with an OA diagnosis may to some extent reflect the misclassification effects identified in Chapter Four, section 4.3.3 (partial over-inclusion of clinical OA through use of joint pain codes that in fact do not always represent OA). The possibility that patients with joint pain and OA are a homogenous population was discussed by Jordan et al.,[17] concluding that the populations of older adults (50 years and over) with joint pain and OA are not in fact homogenous. An OA diagnosis may represent a group more akin to a stereotypical OA patient (older, more likely to be overweight, more symptomatic) than people with joint pain.[17]

The highest level of morbidity (BNF chapter count 10+) was associated with reduced odds of several template-derived quality measures including pain and function assessment, topical NSAID assessment of treatment, education/information provision, and exercise advice. The association between recorded paracetamol prescription and level of total morbidity as measured by a count of BNF chapters could have been to a small extent self-fulfilling – the prescribed drugs for OA (if prescribed in the 12 months before the index consultation) adding to the apparent level of morbidity as measured. However, reduced odds of topical NSAID prescription was seen in the highest morbidity band. The highest morbidity burden being negatively associated with several assessment and advice indicators suggests that in patients with higher apparent morbidity loads, either clinicians or patients have a preference for further pharmacological rather than non-pharmacological management.

It has previously been shown that a decision to use X-ray investigation at a consultation is associated with other aspects of care for knee pain, including increased odds of referral to physiotherapy, rheumatology or orthopædics and reduced odds of advice regarding exercise, even before a result is available.[12] In the adjusted multilevel analyses, recorded X-ray use *before* the index consultation in phase two was not significantly associated with any of the template-derived indicators or routinely-recorded prescribing or referral patterns. There was a significant association between a recorded X-ray within 14 days *after* a clinical OA consultation and achievement of many of the template-derived indicators (pain, function, consideration of paracetamol or topical NSAIDs,

information provision and exercise advice) but not the routinely-recorded prescribing or referral indicators. Given that associations with a recorded X-ray after a clinical OA consultation had been adjusted for all other covariates (including recorded X-ray prior to the index consultation, total morbidity load, and multiple clinical OA consultations) one possible explanation for the association between X-ray use after a clinical OA consultation and achievement of the associated indicators is that some clinicians requested X-rays as a component of what might be termed a 'thorough OA consultation.' The reason for this disparity is unclear but may reflect unmeasured characteristics such as social class or disease severity.

Outcomes for surgical procedures have been found to be related to the volumes of the procedures undertaken at a provider and surgeon level.[290-292] It is conceivable that the increased achievement of quality indicators by clinicians who undertake more than the median number of index consultations was a similar phenomenon, that their greater clinical OA workload was associated with higher recorded quality. On the other hand, it is possible that this represents information bias, in that clinicians who achieved lower levels of recorded quality of care may have seen as many patients with clinical OA as their higher-achieving counterparts but not recorded diagnostic information about them in a coded way (and so for the purposes of an EHR review based on Read codes, those patients would not form a part of the numerator or denominator). However, it seems unlikely that there was a major problem with ascertainment bias as odds of OA diagnosis rather than joint pain coding (Chapter Four) were unrelated to the staff index consultation count, suggesting that musculoskeletal clinical workload did not affect coding behaviour, and the recorded prevalence of total clinical OA was broadly stable between periods, as described fully in Chapter Seven, section 7.3.1, and in keeping with national and international comparisons of OA consultation prevalence.

There were significant differences between practices overall in the quality of care for OA as measured through the template-derived, pharmacological management and referral processes. This type of variation would be consistent with the variability in QOF scores between practices such as reported by Ashworth.[293]

Comparison with previous literature

Broadbent et al.[69] found that female sex and older age were negatively associated with provision of information, and older age was positively associated with an initial trial of paracetamol, and with ibuprofen or a COX-2 selective NSAID where an oral NSAID was used. In Broadbent's study, more severe OA (defined as failure to respond to therapy) was positively associated with assessment for function and pain within the prior 12 months. Severity was not measured in the MOSAICS medical record review study, but neither age nor sex showed consistent effects across quality achievement or recorded clinical management in this study. Li et al.[80] found that women were more likely to receive a record of weight loss advice than men (OR 2.64), as were people aged 55-64 compared to those aged 75 or over (OR 1.96). Li et al. also identified that neither a record of exercise advice nor assessment for assistive devices had a significant association with age or sex.

6.5.2 Limitations

Many of the limitations applicable elsewhere were also relevant to this chapter, including the issues of multiple comparison, dataset limitations (Chapter Four, section 4.4.3 and Chapter Five, section 5.4.2), and interpretation of the VPC (Chapter Four, section 4.4.3).

The template did not always trigger as designed, and a small group of patients who should have formed part of the denominator for this study were excluded. There was no reason to believe that this forms a deliberate selection bias as there was no facility for clinicians to prevent the template triggering, and any 'escape' from the template should still have flagged the record as the template having triggered. The age/sex distribution for people in the template triggering and non-triggering categories were broadly similar, and all practices were affected by the template not triggering as anticipated with a range of 1-18% of patients consulting with clinical OA. The differences in successful triggering between practices were unexplained. The only reason that a bias would have had an effect on the results is if there was a difference in characteristics between patients for whom the template was or was not triggered. Before implementation of the template in routine care, the

reasons behind the failure to trigger consistently for all specified Read codes should be identified and resolved.

6.5.3 Conclusion

Overall, the template-derived data was found to have added substantially to the information available to assess the quality of care for OA. The next Chapter will assess whether the introduction of the template was associated with improvements in routinely-recorded indicators of care.

Chapter Seven: Effects of the template on disease coding and management of osteoarthritis: a before-and-after study

7.1 Introduction

As noted in Chapter Three, the consultation recording template was intended to be an aid to recording of consultations for clinical OA in the MOSAICS study rather than an intervention to promote adherence to the NICE guidance in its own right. Nevertheless, given the evidence base for templates in changing clinical behaviour, [152,162] it was necessary to determine the extent to which the template was associated with a change in clinical behaviour in the study.

The baseline assessment of recorded quality of care for OA (Chapter Five) has shown that various quality indicators were not routinely captured by the medical record, including pain and functional assessment, provision of education or information regarding OA, weight loss, and exercise advice. Chapter Six reported the quality of care as recorded through the template, with a comparison between the template-derived measures and their routinely-recorded counterparts where possible. However, it is not possible to use the indicators recorded through the template to assess any effect of the template itself on delivery of quality care. Rather, such an assessment must depend on indicators that were routinely captured such as those relating to prescriptions, referrals, and weight recording (though the latter could be captured both through the routine record and the template).

The objective of this chapter was to estimate the effects of the template intervention on those aspects of care for OA captured routinely by the medical record.

7.2 Method

The analysis used in this chapter was based upon that determined and undertaken by Professor Jordan for the associated publication *Quality of care for OA: the effect of a point-of-care consultation recording template*[99] (Appendix B), for which the candidate was lead author and led the design and interpretation. This chapter used an extended set of covariates (as used in Chapter

Four to Chapter Six) compared to the analysis for the publication, and is based on a slightly smaller denominator population of people consulting with clinical OA as defined in Chapter Four (diagnosed OA or peripheral joint pain, excluding shoulder or elbow disease). This analysis was devised and performed by the candidate.

As noted in previous chapters, the MOSAICS study was divided into periods of six months each. The first two 6-month periods represented the 12 months immediately before template installation and the third period covered the six months immediately after template installation in all eight practices (see Figure 1-1, page 20). Although there is no evidence of a seasonal variation in the processes of care in general practice,[294] a seasonal fluctuation in practice nurse workload has previously been identified[295] which could have an impact on the delivery of some chronic disease management. Therefore, period one was used as the calendar control to which recording of management in period three was compared. Management in period two was also compared to that in period one to assess whether there was evidence of any natural change in management over time. As in previous analyses, patients were allocated to an index clinician (the first clinician to make a diagnosis of OA in a period, or, if no OA diagnosis was made in a period, the first clinician to record a relevant joint pain code). To link each aspect of care in time to a consultation for clinical OA, a period of 14 days from each such consultation was permitted for the care to be recorded, as previously used by the RIPCHS and consistent with analyses in previous chapters.[99,264]

It was considered possible that the introduction of the template could have caused clinicians to either avoid recording a clinical OA diagnosis (to avoid the template triggering during a consultation) or, potentially, to have caused clinicians to become habituated to recording and managing clinical OA (and thereby record more cases). Therefore, the analysis first considered the diagnostic coding behaviours of clinicians (recording of OA or joint pain in patients within periods), to determine whether or not rates of clinical OA morbidity recording were stable across periods.

To determine the effect of the template on recorded management of clinical OA, indicators of clinical assessment and management (from the systematic review,[107] NICE guidance,[5] or British

National Formulary[217]) described in the baseline quality assessment (Chapter Five) were used to assess the change in quality of care received by populations of patients consulting in each period. However, as the data used had to rely on the routinely recorded (non-template data), certain data were less feasible to use (such as a record of pain and function assessment, provision of education or information, exercise or weight loss advice, risk assessment for oral NSAIDs, and referral for non-ambulatory aids and devices assessment) due to very low levels of routinely recorded indicator achievement in the first two six-month periods and so were excluded from this analysis. Some pharmacological interventions were also not suitable for comparison of care between periods (capsaicin, glucosamine, etoricoxib 60mg, weight loss agents, and rubefacients) as the numerators for these interventions were too small to make meaningful comparisons. Comparisons of OA management between periods were therefore based on data representing weight recording, pharmacological management (paracetamol, topical NSAIDs, opioids, oral NSAIDs, gastroprotection with PPIs), recorded X-ray usage, and referrals (physiotherapy, or all onward referral), using the previously described definitions. Of the feasible routinely recorded management options, those considered most likely to be influenced by the template (as included as template domains) were weight recording, paracetamol and topical NSAID prescribing, and physiotherapy referral. The remaining indicators examined were expected to remain relatively stable as they did not form part of the template recording prompt.

Table 7-1: Routinely recorded measures used to estimate the effect of the template on recorded quality of care for OA

Indicator domain	Indicator origin	Assessment data source(s)	Associated template domain	Anticipated change if template functioned as an intervention
Weight/BMI record	[Prerequisite for advice indicator]	Routinely recorded or template	Yes	Increase
Prescription for paracetamol	Systematic review	Routinely recorded prescriptions	Yes	Increase
Prescription for topical NSAIDs	NICE guidance	Routinely recorded prescriptions	Yes	Increase
Exercise or physiotherapy referral	Systematic review	Routinely recorded codes	Yes	Increase
Prescription for opioid analgesics	NICE guidance	Routinely recorded prescriptions	No	No effect or possible reduction as paracetamol/topical NSAID increases
Prescription for oral NSAIDs	NICE guidance	Routinely recorded prescriptions	No	No effect or possible reduction as paracetamol/topical NSAID increases
Prescription for oral NSAIDs in people with recorded comorbidity as a relative contraindication	NICE guidance	Routinely recorded prescriptions	No	No effect or possible reduction as paracetamol/topical NSAID increases
Co-prescription of gastroprotection (for people prescribed oral NSAIDs)	Systematic review	Routinely recorded codes	No	No effect
Recorded use of X-rays within 14 days of a consultation for peripheral joint clinical OA	NICE guidance	Routinely recorded codes	No	No effect
Any onward referral	[Aggregation of NICE possible referral interventions]	Routinely recorded codes	No	No effect

7.2.1 Statistical analysis

The denominator population was assumed to be static, as discussed in Chapter Four, and the practice population used was the same as that used for the baseline periods one and two.

Age-standardised period consultation prevalence rates per 100,000 population were calculated for each practice and overall as described in Chapter Four (section 4.2.4), using the European Standard Population.[226] The differences in rates between periods were calculated via the standardised rate ratio and its 95% CI.[296]

The routinely recorded care for clinical OA in the populations consulting in the three six-month periods of interest were described initially. The frequency of achievement by month was also plotted across periods one to three as used in the associated publication[99] and are reproduced here.

Recorded management in periods two and three (the six month periods immediately before and after template installation) were then compared to that in period one using two-level binary logistic multilevel models (patients within index clinicians). The model was adjusted for the same variables used in previous chapters: sex, age, site of disease, BMI status, morbidity burden as assessed by a count of drug types, single or multiple clinical OA consultations within periods, recorded X-ray use prior to the index consultation, the number of index consultations undertaken by the staff member consulted (dichotomised into above or at-or-below the median number), and the practice.

Patients' consultations within periods were aggregated to a single summary of the diagnosis and management within that period. Consultations by the same patient in different time periods were treated independently, as attempts to use a three-level model to account for repeated consultations within patients across periods were unsuccessful.

As well as the primary analysis, which included all clinical OA consulters in each of the three periods, sensitivity analyses were undertaken. The effect of the template was estimated using the same methodology as noted above in patients for whom the template was recorded as having triggered in period three (*n*=1724) (but with no change to the eligible population in periods one and two), and in those with a recorded template entry other than weight or BMI (*n*=1142). Whereas the main

analysis estimates the effect of the template implementation strategy (akin to an intention-to-treat analysis), the first of the two sensitivity analyses is regarded as a truer estimate of the effect of the template itself whilst the second sensitivity analysis estimates the effect in people in whom clinicians have engaged with the template use.

The data were analysed for the descriptive epidemiology in Microsoft® Excel®,[231] CIA,[232] StatsDirect,[297] and SPSS v21[230] and, for the multilevel modelling, in Stata 14.1[298] and MLwiN[234] using the runMLwiN[235] command.

7.3 Results

7.3.1 Consultations by period

Over the three periods, 4393 unique patients consulted for clinical OA: 3496 in just one period (79.6%), 771 in two periods (17.6%), and 126 (2.9%) in all three periods. There was a slight trend to increasing frequency of consultation with 1750 patients consulting in period one, 1821 in period two, and 1845 (of whom 1724 triggered the template) in period three (approximately a 5% overall increase). The practice-specific age-standardised rates are shown in Table 7-3, with the SRRs and their 95% CI.

Table 7-2: Patients consulting by time period with OA or joint pain classification

	Patients consulting within each time period		
	1: 6-12 months before template fire *n* (%)	2: 0-6 months before template fire *n* (%)	3: 0-6 months after template fire *n* (%)
Joint pain	1174 (67)	1184 (65)	1237 (67)
OA	576 (33)	637 (35)	608 (33)
Total	1750	1821	1845

There was substantial variation within practices between periods in the age-standardised rates of joint pain and OA between periods. There was evidence only of a statistically significant increase in recorded non-OA joint pain and combined clinical OA between periods two and three in practice six and an increase in combined clinical OA in practice eight between periods one and three, and two and three.

Table 7-3: Age-standardised disease rates (ASR) per 10,000 adults aged ≥45 years with 95% CI, and standardised rate ratios (SRR) with 95% CI, total and by practice

			Practice 1	Practice 2	Practice 3	Practice 4	Practice 5	Practice 6	Practice 7	Practice 8	Total
Joint pain	**ASR**	**Period 1**	287 (256,319)	229 (164,294)	382 (325,440)	340 (288,391)	528 (451,604)	390 (305,474)	438 (368,509)	229 (163,295)	230 (217,243)
		Period 2	309 (276,342)	253 (183,322)	366 (309,422)	383 (328,437)	486 (413,560)	300 (227,374)	440 (370,511)	219 (153,284)	232 (219,246)
		Period 3	291 (260,323)	217 (153,282)	419 (358,480)	348 (296,400)	542 (465,620)	429 (339,518)	467 (394,539)	320 (241,399)	242 (228,255)
	SRR	**Period 2:1**	1.07 (0.92,1.25)	1.10 (0.74,1.64)	0.96 (0.77,1.19)	1.13 (0.91,1.39)	0.92 (0.75,1.14)	0.77 (0.56,1.07)	1.01 (0.80,1.26)	0.95 (0.63,1.44)	1.01 (0.93,1.10)
		Period 3:1	1.01 (0.87,1.18)	0.95 (0.63,1.43)	1.10 (0.89,1.35)	1.03 (0.83,1.27)	1.03 (0.84,1.26)	1.10 (0.81,1.49)	1.06 (0.85,1.33)	1.40 (0.96,2.04)	1.05 (0.97,1.14)
		Period 3:2	0.94 (0.81,1.10)	0.86 (0.57,1.29)	1.15 (0.93,1.42)	0.91 (0.74,1.12)	1.12 (0.91,1.37)	**1.42 (1.03,1.97)**	1.06 (0.85,1.32)	1.46 (1.00,2.15)	1.04 (0.96,1.13)
Formally diagnosed OA	**ASR**	**Period 1**	273 (242,305)	162 (104,219)	114 (82,145)	111 (82,140)	171 (127,215)	64 (29,97)	77 (47,107)	202 (140,263)	113 (104,122)
		Period 2	292 (260,325)	178 (119,238)	137 (103,171)	150 (116,183)	197 (150,243)	34 (9,59)	87 (55,119)	216 (152,280)	125 (115,135)
		Period 3	267 (237,298)	251 (180,322)	134 (99,168)	139 (106,171)	156 (115,215)	32 (8,56)	53 (28,78)	274 (201,347)	118 (109,128)
	SRR	**Period 2:1**	1.07 (0.91,1.25)	1.10 (0.68,1.79)	1.21 (0.83,1.75)	1.35 (0.95,1.90)	1.15 (0.81,1.63)	0.54 (0.22,1.34)	1.13 (0.66,1.92)	1.07 (0.70,1.64)	1.11 (0.99,1.24)
		Period 3:1	0.98 (0.83,1.15)	1.55 (0.99,2.44)	1.18 (0.80,1.72)	1.25 (0.88,1.77)	0.91 (0.63,1.32)	0.51 (0.20,1.27)	0.69 (0.38,1.26)	1.36 (0.90,2.04)	1.05 (0.93,1.17)
		Period 3:2	0.92 (0.78,1.07)	1.41 (0.91,2.18)	1.10 (0.92,1.32)	0.93 (0.68,1.28)	0.80 (0.56,1.14)	0.94 (0.33,2.69)	0.61 (0.34,1.10)	1.27 (0.85,1.89)	0.94 (0.85,1.06)
Combined clinical OA	**ASR**	**Period 1**	561 (516,605)	391 (304,478)	496 (430,562)	451 (392,510)	699 (610,787)	452 (361,544)	515 (438,592)	431 (339,523)	343 (327,359)
		Period 2	601 (555,647)	431 (340,523)	503 (437,569)	532 (468,597)	683 (596,770)	334 (256,412)	527 (450,604)	435 (341,528)	357 (341,374)
		Period 3	559 (515,605)	468 (373,564)	553 (483,623)	487 (426,548)	699 (610,787)	461 (368,553)	520 (443,596)	593 (484,703)	360 (343,377)
	SRR	**Period 2:1**	1.07 (0.96,1.20)	1.10 (0.81,1.50)	1.01 (0.84,1.22)	1.18 (0.99,1.41)	0.98 (0.82,1.17)	0.74 (0.54,1.01)	1.02 (0.83,1.26)	1.01 (0.74,1.37)	1.04 (0.98,1.11)
		Period 3:1	1.00 (0.89,1.11)	1.20 (0.89,1.62)	1.12 (0.92,1.34)	1.08 (0.90,1.29)	1.00 (0.84,1.20)	1.02 (0.77,1.35)	1.01 (0.82,1.24)	**1.38 (1.04,1.83)**	1.05 (0.98,1.12)
		Period 3:2	0.93 (0.83,1.04)	1.09 (0.81,1.46)	1.10 (0.92,1.32)	0.91 (0.77,1.09)	1.02 (0.86,1.22)	**1.38 (1.01,1.87)**	0.99 (0.80,2.21)	**1.37 (1.03,1.81)**	1.01 (0.94,1.08)

Statistically significant standardised rate ratios are shown in **bold**

7.3.2 Descriptive epidemiology

Characteristics of patients consulting within each time period are shown in Table 7-4. Most of the characteristics remain fairly stable in prevalence, although there was evidence of better BMI status recording (21.3% patients in the 'unknown' category in period three compared to 35.7% in period one).

The levels of recording of the outcome measures from the routinely-collected data sources are shown in Table 7-5. These measures were relatively stable across periods for most measures, though there was an apparent increase in weight recording (8.9% in period one to 23.4% in period three) and prescription of paracetamol or topical NSAIDs (25.5% in period one to 35.2% in period three), within 14 days of a consultation for clinical OA.

The change in weight recording and prescription of relatively safe pharmaceutical options can be seen as shown through the charted monthly data in Figure 7-1.[99] The timeline commenced 12 months before template introduction (i.e. at the start of period one). By inspection, an increase in weight recording, and paracetamol and topical NSAID prescribing can be seen at the start of period three (template introduction), though weight recording does appear to fall back somewhat though remaining above baseline levels as seen in periods one and two.

Table 7-4: Patient characteristics within each time period

		Period 1 *n* (%)	Period 2 *n* (%)	Period 3 *n* (%)
	Total	1750	1821	1845
Sex	Male	718 (41.0)	693 (38.1)	760 (41.2)
	Female	1032 (59.0)	1128 (61.9)	1085 (58.8)
Age-band	45-64	812 (47.4)	850 (47.4)	880 (48.5)
	65-74	475 (27.7)	482 (26.9)	469 (25.8)
	75-84	336 (19.6)	354 (19.8)	370 (20.4)
	85+	91 (5.3)	106 (5.9)	97 (5.3)
Site affected	Knee	904 (51.7)	881 (48.4)	928 (50.3)
	Hip	376 (21.5)	414 (22.7)	378 (20.5)
	Ankle/foot	116 (6.6)	113 (6.2)	130 (7.0)
	Hand/wrist	148 (8.5)	172 (9.4)	168 (9.1)
	Unspecified	91 (5.2)	118 (6.5)	122 (6.6)
	Multiple	115 (6.6)	123 (6.8)	119 (6.4)
BMI status	BMI <25	235 (13.4)	276 (15.2)	327 (17.7)
	BMI 25 to <30	418 (23.9)	421 (23.1)	548 (29.7)
	BMI 30+	473 (27.0)	586 (32.2)	577 (31.3)
	BMI unknown	624 (35.7)	538 (29.5)	393 (21.3)
Count of BNF chapters in 12m prior	0-4	572 (32.7)	547 (30.0)	560 (30.4)
	5-9	536 (30.6)	623 (34.2)	614 (33.3)
	10+	642 (36.7)	651 (35.7)	671 (36.4)
Multiple consultations in period	Single	1171 (66.9)	1228 (67.4)	1294 (70.1)
	Multiple	579 (33.1)	593 (32.6)	551 (29.9)
XR recorded within 21 months prior to index consultation	No	1652 (94.4)	1714 (94.1)	1758 (95.3)
	Yes	98 (5.6)	107 (5.9)	87 (4.7)
Consultation with staff member holding greater or fewer than the median number of index consultations	Below the median	208 (11.9)	137 (7.5)	197 (10.7)
	Above the median	1542 (88.1)	1684 (92.5)	1648 (89.3)
Practice	1	619 (35.4)	660 (36.2)	622 (33.7)
	2	80 (4.6)	87 (4.8)	94 (5.1)
	3	223 (12.7)	227 (12.5)	245 (13.3)
	4	228 (13.0)	268 (14.7)	247 (13.4)
	5	244 (13.9)	239 (13.1)	245 (13.3)
	6	95 (5.4)	72 (4.0)	97 (5.3)
	7	176 (10.1)	183 (10.0)	180 (9.8)
	8	85 (4.9)	85 (4.7)	115 (6.2)

Table 7-5: Clinical process measures for the care of OA derived from routinely-recorded data, by time period

			Period 1 *n* (%)	Period 2 *n* (%)	Period 3 *n* (%)
		Total	1750	1821	1845
Relevant template domain	Any OA or only joint pain (JP) within period	JP	1174 (67.1)	1184 (65.0)	1237 (67.0)
		OA	576 (32.9)	637 (35.0)	608 (33.0)
	Weight record within 14 days of a consultation for clinical OA	No	1595 (91.1)	1654 (90.8)	1413 (76.6)
		Yes	155 (8.9)	167 (9.2)	432 (23.4)
	Paracetamol prescription within 14 days of a consultation for clinical OA	No	1517 (86.7)	1591 (87.4)	1529 (82.9)
		Yes	233 (13.3)	230 (12.6)	316 (17.1)
	Topical NSAID prescription within 14 days of a consultation for clinical OA	No	1482 (84.7)	1547 (85.0)	1387 (75.2)
		Yes	268 (15.3)	274 (15.0)	458 (24.8)
	Paracetamol or a topical NSAID within 14 days of a consultation for clinical OA	No	1303 (74.5)	1362 (74.8)	1196 (64.8)
		Yes	447 (25.5)	459 (25.2)	649 (35.2)
	Physio referral within 14 days of a consultation for clinical OA	No	1641 (93.8)	1717 (94.3)	1720 (93.2)
		Yes	109 (6.2)	104 (5.7)	125 (6.8)
No relevant template domain	Opioid prescription within 14 days of a consultation for clinical OA	No	1179 (67.4)	1222 (67.1)	1259 (68.2)
		Yes	571 (32.6)	599 (32.9)	586 (31.8)
	Oral NSAID prescription within 14 days of a consultation for clinical OA	No	1443 (82.5)	1525 (83.7)	1547 (83.8)
		Yes	307 (17.5)	296 (16.3)	298 (16.2)
	Oral NSAID prescribed where comorbidities as a relative contraindication are recorded (n=2639)	No	730 (86.1)	787 (88.3)	781 (86.8)
		Yes	118 (13.9)	104 (11.7)	119 (13.2)
	PPI prescription within 14 days of a consultation for clinical OA	No	1425 (81.4)	1430 (78.5)	1488 (80.7)
		Yes	325 (18.6)	391 (21.5)	357 (19.3)
	X-ray investigation within 14 days of a consultation for clinical OA	No	1468 (83.9)	1511 (83.0)	1573 (85.3)
		Yes	282 (16.1)	310 (17.0)	272 (14.7)
	Referral within 14 days of a consultation for clinical OA	No	1352 (77.3)	1408 (77.3)	1474 (79.9)
		Yes	398 (22.7)	413 (22.7)	371 (20.1)

7.3.3 The effect of time period on outcomes: multilevel model

There was no evidence of a substantial change in any of the management options between periods one and two (no naturally-occurring temporal change in OA quality of care), and none of the ORs were found to be statistically significant. Of the management aspects with a relevant template domain, the OR of a weight record in period three compared to period one was 3.06 (95% CI 2.46,3.82), with prescriptions of paracetamol (adjusted OR 1.46, 95% CI 1.19,1.79) and topical NSAIDs (adjusted OR 2.00, 95% CI 1.64,2.44) also increasing. When the paracetamol and topical

NSAID options were combined, there was a significant increased odds of prescription of either option (adjusted OR 1.79 in period three, 95% CI 1.52,2.12). There was no significant change physiotherapy referral (adjusted OR 1.10, 95% CI 0.82,1.48). Other aspects of care showed no significant change (Table 7-6).

Figure 7-1: Recorded assessment, prescriptions and referral within 14 days of a consultation for clinical OA, by month, in phases one and two

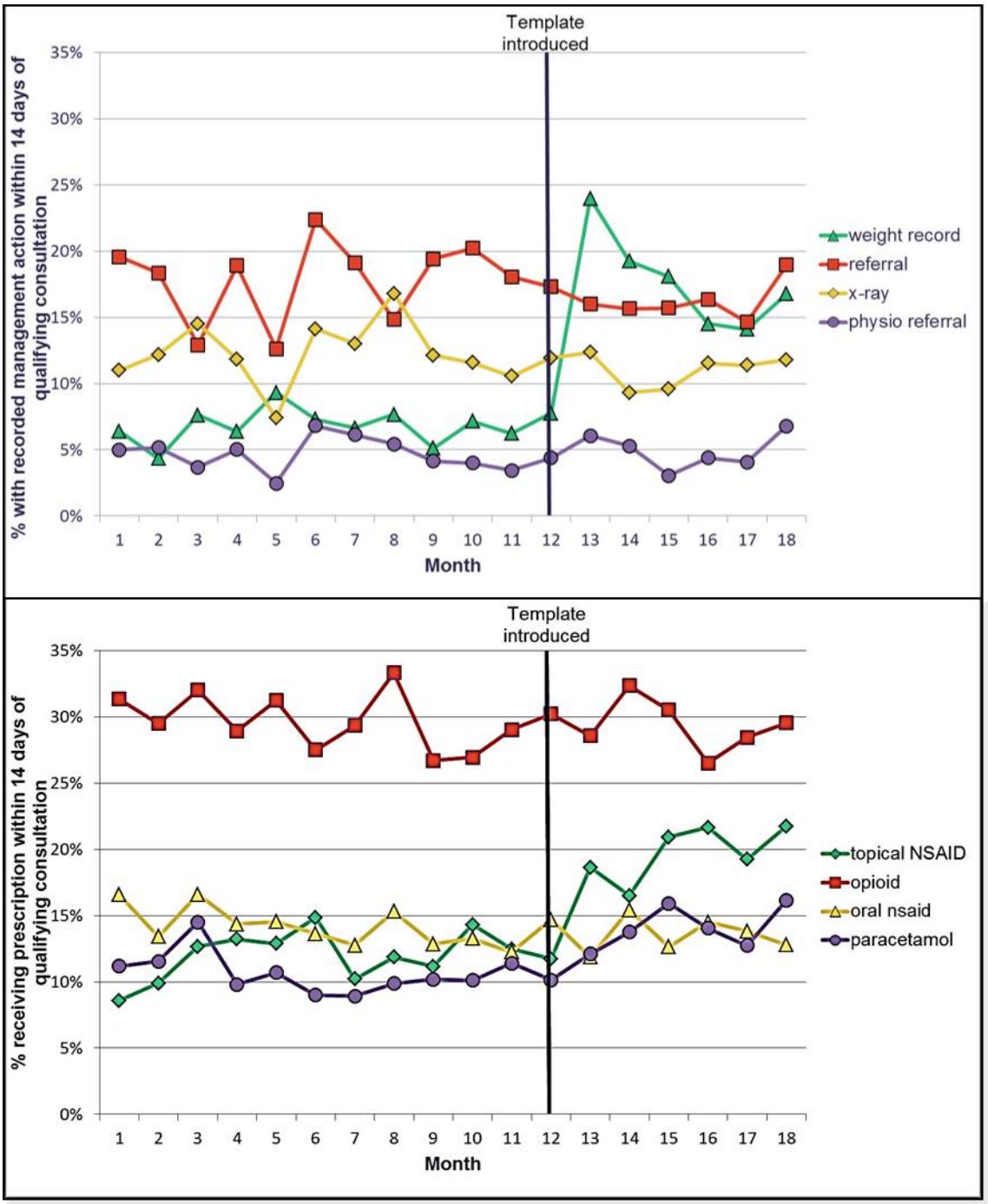


Table 7-6: Associations between routinely recorded care for OA and time period

		Time period (vs. Period 1)			
		Period 2		Period 3	
		Unadjusted OR (95% CI)	Adjusted[a] OR (95% CI)	Unadjusted OR (95% CI)	Adjusted[a] OR (95% CI)
	Diagnosed OA vs. joint pain	1.04 (0.89,1.21)	1.04 (0.88,1.23)	1.01 (0.87,1.19)	1.01 (0.85,1.21)
Relevant template domain	Weight record	1.02 (0.80,1.30)	0.95 (0.74,1.22)	↑ **3.22 (2.60,3.98)**	↑ **3.06 (2.46,3.82)**
	Paracetamol prescription	0.94 (0.77,1.15)	0.92 (0.74,1.14)	↑ **1.39 (1.15,1.69)**	↑ **1.46 (1.19,1.79)**
	Topical NSAID prescription	0.97 (0.80,1.17)	1.00 (0.81,1.23)	↑ **1.81 (1.51,2.17)**	↑ **2.00 (1.64,2.44)**
	Paracetamol or topical NSAID prescription combined	0.98 (0.84,1.15)	0.99 (0.83,1.17)	↑ **1.61 (1.38,1.88)**	↑ **1.79 (1.52,2.12)**
	Physiotherapy referral	0.90 (0.66,1.22)	0.87 (0.64,1.17)	1.13 (0.83,1.52)	1.10 (0.82,1.48)
No relevant template domain	Opioid prescription	1.03 (0.89,1.19)	1.01 (0.86,1.18)	0.97 (0.84,1.12)	0.98 (0.84,1.15)
	Oral NSAID prescription	0.91 (0.76,1.10)	0.93 (0.61,1.12)	0.89 (0.73,1.07)	0.94 (0.78,1.15)
	Oral NSAIDs in the presence of contraindicating comorbidity	0.79 (0.59,1.06)	0.83 (0.61,1.12)	0.92 (0.69,1.24)	0.99 (0.74,1.34)
	PPI with oral NSAIDs	1.11 (0.78,1.57)	1.27 (0.86,1.87)	0.98 (0.69,1.40)	1.07 (0.72,1.58)
	Recorded X-ray use	1.07 (0.87,1.31)	1.11 (0.90,1.38)	0.92 (0.74,1.14)	0.96 (0.77,1.20)
	Any referral	0.99 (0.84,1.17)	0.98 (0.82,1.16)	0.88 (0.74,1.04)	0.87 (0.73,1.04)

[a]adjusted for all covariates

Table 7-7: Associations between routinely recorded care for OA and time period with restriction to patients in period three for whom the template triggered and those for whom at least one template entry was made

		Time period (vs. Period 1)		Time period (vs. Period 1)	
		Period 2	Patients triggering template only in Period 3	Period 2	Patients with at least one template entry in Period 3
		Adjusted[a] OR (95% CI)	Adjusted[a] OR (95% CI)	Adjusted[a] OR (95% CI)	Adjusted[a] OR (95% CI)
	Diagnosed OA vs. joint pain	1.04 (0.88,1.23)	1.03 (0.86,1.23)	1.04 (0.88,1.23)	1.04 (0.85,1.27)
Relevant template domain	Weight record	0.95 (0.74,1.22)	↑ **3.20 (2.56,3.99)**	0.92 (0.72,1.18)	↑ **4.60 (3.64,5.82)**
	Paracetamol prescription	0.93 (0.75,1.15)	↑ **1.46 (1.18,1.79)**	0.93 (0.75,1.15)	↑ **1.53 (1.22,1.92)**
	Topical NSAID prescription	1.00 (0.82,1.23)	↑ **2.06 (1.68,2.51)**	1.01 (0.82,1.24)	↑ **2.96 (2.38,3.68)**
	Paracetamol or topical NSAID prescription combined	0.99 (0.83,1.17)	↑ **1.81 (1.53,2.15)**	1.00 (0.84,1.18)	↑ **2.38 (1.97,2.87)**
	Physiotherapy referral	0.86 (0.64,1.17)	1.13 (0.84,1.53)	[b]0.87 (0.64,1.17)	[b]1.27 (0.94,1.73)
No relevant template domain	Opioid prescription	1.01 (0.87,1.19)	0.99 (0.84,1.16)	1.01 (0.86,1.19)	0.95 (0.79,1.14)
	Oral NSAID prescription	0.93 (0.77,1.13)	0.95 (0.78,1.16)	0.93 (0.77,1.13)	0.98 (0.79,1.22)
	Oral NSAIDs in the presence of contraindicating comorbidity	0.83 (0.61,1.12)	0.99 (0.74,1.34)	0.82 (0.60,1.11)	1.06 (0.76,1.48)
	PPI with oral NSAIDs	1.27 (0.86,1.87)	1.07 (0.72,1.58)	1.27 (0.85,1.87)	1.12 (0.72,1.75)
	Recorded X-ray use	1.11 (0.90,1.37)	0.97 (0.78,1.21)	1.11 (0.90,1.38)	1.13 (0.89,1.43)
	Any referral	0.98 (0.83,1.16)	0.89 (0.75,1.07)	0.98 (0.83,1.16)	0.89 (0.73,1.08)

[a]adjusted for all covariates; [b]PQL1 estimation, patients within clinicians

A repeat analysis of the same outcomes, but restricting the eligible population in period three to only those in whom the template triggered, showed very little change to the OR estimates, as shown in Table 7-7. A further restriction to patients in period three who had at least one template entry recorded suggested a stronger association between the template with weight recording and relatively safe pharmacological management options.

7.4 Discussion

7.4.1 Main findings and comparison with previous literature

Consultation rates for clinical OA were fairly stable across the three periods, though there was some variation between practices and overall there was a slight increase in recorded consultation prevalence which did not reach statistical significance. There was no evidence of any significant change in the routinely recorded management of OA between periods one and two (i.e. the periods prior to template introduction), suggesting that over this period there was no 'naturally-occurring' quality improvement in OA care. The template installation was associated with improvements in weight recording and use of relatively safe pharmacological management options (paracetamol and topical NSAIDs) between periods one and three (the calendar control period and the first six months' template use). There was no statistically significant change in any of the other management options. Sensitivity analyses to distinguish effects in sub-populations of consulters in period three (those in whom the template actually triggered and those in whom at least one entry was made) showed some indication of a greater effect in patients with at least one entry.

Rose et al. reported that recording templates in a USA study were associated with improvements in documentation (of the clinical, family, and social history, and of clinical examination findings) but they found no effect on coding.[299]

Evidence from the Cochrane review of the effects of computer reminders,[152] described in Chapter Three, section 3.3, identified a median 4.2% (interquartile range [IQR] 0.8-18.8%) improvement in process measures overall, with recommended medication ordering improving by a median 6.2% (IQR 3.0-28.0%). In this study, the weight recording process improved from around 9% in periods

one and two to 23.4% in period three. The improvement in weight recording was rather more than would have been expected by the Cochrane review evidence. It was uncertain to what extent the increased recording of weight reflected an increase in the number of patients actually weighed rather than just an increase in the proportion of the measurements made being recorded.

Paracetamol and topical NSAIDs prescribing processes increased from around 13-15% in periods one and two to 17.1-24.8% in period three, in line with what would be expected from the Cochrane review. The increase in prescribing was most likely to reflect an actual change in clinical behaviour, since prescriptions, with few exceptions, are electronically generated and thereby recorded in a standard manner in the clinical computer systems.

Although there was a trend toward increased physiotherapy referral, this was small and did not achieve statistical significance. The absence of any significant change in physiotherapy referral suggested that the template was not effective at changing clinical behaviour in this respect. This may have been due to other influences such as beliefs about the benefits of physiotherapy for OA[288] (at either the patient or clinician level), or access to physiotherapy services (either due to lack of sufficient service provision,[289] thereby discouraging referral for conditions considered to be either less important or suitable for self-management, or due to open access without referral such as through Physio Direct[169]). Also, the physiotherapy prompt was the final item on the template, and it is possible that there was an order effect such that the earlier items received preferential attention, though no evidence regarding the order of items in reminder systems has been identified in the literature.

The lack of any statistically significant changes in the other routinely recorded clinical management, excluded from and therefore not directly influenced by the template, was notable.

Although some improvements in recorded quality of care were identified in this analysis, it was not clear from the available information how durable the changes would be in routine clinical practice. It was unclear from the prior literature how long the effect of the template may be expected to last.[162] This warrants further investigation.

7.4.2 Limitations

There were some limitations to this study. Limitations relating to the interpretation of multiple comparisons, as described in Chapter Five, were applicable to this chapter. The linkage of processes of care to clinical OA consultations through a 14-day timeframe should not have caused bias in the assessment of the template effect.

All observational studies suffer from limitations in attribution of causality to an intervention.[300] However, this study was based upon all consulters for clinical OA in period three and so risks of selection and information biases had been ameliorated. There remained a risk that the association between improvements in recorded quality of care and the template introduction was not due to the template but rather to another cause, such as study inception overall. That is, the association between the template as an exposure and the apparent quality improvement may be the result of confounding, such as by clinician self-directed learning resulting from inclusion into an academic study, independently related to both the exposure and the outcome. It is possible that the commencement of the study in the eight practices prompted greater awareness of OA management in general.

Given the temporal sequence and strength of association between several of the domains included in the template and improvement in quality measures, not seen in any of the non-template domains, it is considered likely that it was the template itself that prompted the quality improvement.

The fact that this analysis does not account for repeated consultations by patients across periods was an analytical limitation. This may have caused some underestimation of standard errors of the coefficients in the multilevel models as the clustering of clinical management within patients was not accounted for by this analysis.[229] Ideally, an analysis of consultations within patients within clinicians within practices as a four-level model would have been undertaken but this would have dramatically increased the complexity of the model (as patients were able to consult with more than one clinician within practices, a cross-classified model would have been required) and a larger

dataset would have been needed for this degree of complexity to be analysed. The small number of practices (clusters) also limited the potential for the use of a more complex model that would have incorporated practice as the topmost level.

7.4.3 Conclusion

In conclusion, the template seems not to have been a neutral intervention but rather was associated with some important changes in clinical behaviour. Although a causal effect has not been conclusively demonstrated, either in terms of promoting better recording (such as with weight records) or in terms of promoting increased use of relatively safe pharmaceutical options, it seems reasonable to conclude that the template does have a small to moderate effect. It is not clear from the available data why no significant change was seen to the low levels of physiotherapy referral but it is possible that this related to factors not addressed by the template, including practitioner beliefs about exercise in OA management, or service availability.

In Chapter Eight, the use of the quality indicators in a trial of the MOSAICS intervention will be described.

Chapter Eight: Implementation of quality indicators in a cluster-randomised trial of a model OA consultation

8.1 Introduction

It has previously been shown (in Chapter Two) how quality of care for OA can be assessed through evidence-based quality indicators,[107] and that primary care for OA has been found to be suboptimal when compared to quality indicators and guideline adherence. The limited utility of the routinely-recorded medical record data for assessment of quality of care through the quality indicators was described in Chapter Five, and Chapter Six & Chapter Seven described the quality assessments that could be made through medical records enhanced by the template recording system and the effect of the template on recorded care.[99]

There are a few examples of clinical trials that use indicators of care as outcome measures, such as the use of prescribing process indicators in a trial of hospital feedback on care of people with myocardial infarction,[301] a cluster trial of payment for performance incentives on prescribing processes and surrogate patient outcomes,[302] and a trial of the introduction of a system to improve prevention of maternal HIV transmission through HIV testing and prescription of antiretroviral drugs.[303] However, the use of indicators as outcome measures is not a highly-developed trial methodology.

The MOSAICS study,[1] outlined in Chapter One, section 1.7, was conceived as a trial of a complex intervention with the intention of increasing uptake of the NICE-recommended interventions for OA in primary care.[5] The final step in the synthesis and testing of the quality indicators for OA was to implement them in the MOSAICS study, as trial outcome measures. This analysis complements other MOSAICS trial analyses that were not part of this thesis, relating to the individual patient-level clinical effectiveness of the intervention (Dziedzic KS, et al., in submission) and to the survey of the practice populations (Healey E, et al., in submission). This analysis is based on the same data and a similar analysis reported in Jordan KP, et al., in submission, for which the candidate was a co-author and co-designed the analysis, but due to the objectives of this chapter which focus on

exploration of the use and feasibility of quality indicators as outcomes in trials rather than to assess effectiveness of the intervention, used a different approach to adjustment for covariates and undertakes additional analysis relevant to the clustered nature of the trial.

The objective of this chapter is to describe how the effects of a complex intervention to improve uptake of NICE-recommended management strategies for OA in general practice registered populations could be assessed through quality indicators including those recorded via a template integrated in the clinical information system.

8.2 Methods

8.2.1 The MOSAICS trial

This analysis used data taken from the MOSAICS study, the protocol for which has been published previously.[1] Although the objective of this chapter is to describe the use of quality indicators in a trial setting rather than to report the MOSAICS trial in full, the description given here is based upon the CONSORT statement for reporting of cluster trials.[92] The cluster trial took place between May 2012 and February 2014. It was approved by the North West 1 Research Ethics Committee, Cheshire (REC reference: 10/H1017/76, Appendix C) and monitored by an independent Trial Steering Committee and Data Monitoring Committee (trial registration number ISRCTN06984617).

Trial design

MOSAICS was a parallel cluster-randomised controlled clinical trial. A cluster design was considered to be necessary due to a substantial risk of contamination of the control element were clinicians expected to offer either usual care or enhanced care to patients who had been randomised at an individual level. Practices were randomly assigned in a 1:1 ratio to intervention or control arms, stratified by practice size to maintain, as far as possible, a population size balance between arms.

Setting and Population

The trial was undertaken in eight clusters, each a general practice in the North-West Midlands and Cheshire, UK. Practices were resourced for participation in the trial as set out in the trial protocol.[1]

All patients aged 45 years and over, consulting at least once with recorded clinical OA (defined as formally diagnosed OA (excluding shoulder or elbow) or selected hand, hip, knee, or foot joint pain Read codes as described in Chapter Three, section 3.4.3) in the six month period after randomisation and training (trial period four) were eligible for inclusion in this aspect of the trial analysis unless dissent to secondary use of their primary care clinical data had previously been recorded in the medical record, such as has been used by some patients due to privacy concerns.

During the trial baseline period, six months prior to randomisation (see the MOSAICS timeline in Figure 1-1, page 20), practices were encouraged to use the template described earlier in consultations for clinical OA,[99] to capture relevant aspects of OA care not routinely coded in the EHR. The template was automatically triggered by entry of a Read code for clinical OA as defined in the template development (Chapter Three, section 3.4.3). Within the EMIS clinical computer system, it was also possible to manually invoke the template if required (such as if the template had been bypassed in error, or failed to trigger as expected).

Intervention

The development of the trial intervention was not part of this thesis and is described fully elsewhere.[1,304-306] An overview of the intervention is set out below.

The Model OA Consultation

Based on theories of support for behaviour change[307-309] and self-management,[310] and the outcome of Delphi consensus exercises for GPs[305] and nurses,[306] a model OA consultation (MOAC) was developed. This consisted of an initial GP consultation (for all patients consulting with clinical OA, not only people with a new diagnosis) to make and communicate the OA diagnosis, provide a self-management guidebook, and address analgesia requirements, and was followed by a series of up to four consultations with a practice nurse to support self-management of OA.

Training

Practices in the intervention arm received a package of training for the GPs and practice nurses, and trial orientation training for other relevant clinical primary care team members. The GPs received model consultation training, including use of simulated patients, in four sessions totalling seven hours. Practice nurses received training on aspects of self-management including the core NICE interventions (education, physical activity, and maintenance of a healthy weight), goal setting, and pain management (through exercise and analgesia); their training lasted four days and again included simulated patients.

Guidebook

Both GP and practice nurse consultations were designed to be supported by the use of an OA self-management guidebook.[304] This patient-focussed guide is available at http://www.keele.ac.uk/media/keeleuniversity/ri/primarycare/pdfs/OA_Guidebook.pdf [accessed 03.09.2016]. This was designed to be issued at the initial GP consultation and to be used in the interlude between GP and nurse consultations for patients to develop their understanding of OA in advance of the nurse consultation at which any follow up questions could be answered.

Control

Practices in the control arm continued with usual care as in the baseline period (defined as standard GP care supplemented with the template for recording consultations), and with access to written materials except the OA guidebook. No training or resource for additional nurse consultations was provided to the control practices.

Outcomes

The trial outcomes were analysed and reported at two levels: (i) the patient-reported outcomes, in a subgroup of patients consenting to take part in the survey aspect of the trial (Dziedzic KS, et al., in submission), and (ii) the population-level quality of care, reported in Jordan KP, et al. (in submission) and in this analysis. The primary outcomes used here were the rates of achievement

of quality indicators in each study arm, recorded either through the routine record or the template, during the six-month period after randomisation and MOAC training (period four, Figure 1-1). These indicators were those used in the previous chapters, with extension of the analysis to examine in particular rates of recorded provision of written information and written advice regarding exercise and weight loss, as provision of written, as opposed to purely verbal, information and advice was a core component of NICE-recommended OA management.[5] The four outcome domains of quality of care were (i) assessment (pain, function, weight or BMI record, X-ray use), NICE-recommended (ii) core management (information provision, exercise advice, weight loss advice where relevant), (iii) non-pharmacological management (referrals to a physiotherapist), and (iv) pharmacological management (paracetamol, topical NSAIDs, oral NSAIDs, gastroprotection with proton pump inhibitors (PPIs), opioids) (Table 8-1).

Outcomes were measured through codes recorded in the EHR. Just as in earlier chapters, these comprised the indicators identified via routinely recorded management (Chapter Five) and those from the template (Chapter Six and Chapter Seven).

All outcomes were considered to have occurred if they were recorded within 120 days of an index consultation occurring in the trial period (which provided enough time for patients in the intervention arm to have completed nurse clinic follow-up). Measures derived from the routine record also had to be recorded within 14 days of a consultation for clinical OA. For PPIs, the prescription additionally had to have occurred on the same day as a prescription for an oral NSAID.

Table 8-1 gives the direction of change that would be expected at a population level if the primary care for OA was improving. For example, although X-rays are not considered necessary to make a diagnosis of OA nor to determine the need for onward referral,[4] their use may still be appropriate on an individual level; but at a population level, one would expect that better care would be associated with lower aggregated use. One might expect use of relatively safe pharmacological options to increase but for X-rays, oral NSAIDs, and opioids and to decrease if care was consistent with NICE guidance.

Table 8-1: Indicators of quality of care for OA used in the MOSAICS trial

Domain	Quality indicator	Indicator source[a]	Data source	Evidence of achievement (dichotomous outcome)	Change indicative of improved care[b]
Assessment	Pain assessed	Review	Template	Recorded level of pain[c]	Increase
	Function assessed	Review	Template	Recorded level of function[c]	Increase
	BMI measurement/weight record	Review	Template or routine EMR	Recorded BMI or weight	Increase
	X-ray requested	Guideline	Routine EMR	Recorded X-ray of knee, hip, hand, or foot	Decrease
Core interventions	OA information	Review	Template	Recorded as verbal or written; or not appropriate[d]	Increase
	Written OA information	Guideline	Template	Recorded as written	Increase
	Exercise advice	Review	Template	Recorded as verbal or written; or not necessary or not appropriate[d]	Increase
	Written exercise advice	Guideline	Template	Recorded as written	Increase
	Weight loss advice[e]	Review	Template	Recorded as verbal or written; or not appropriate[d]	Increase
	Written weight loss advice	Guideline	Template	Recorded as written	Increase
Non-pharmacological interventions	Consideration of physiotherapy referral	Guideline	Template	Recorded as offered; or not necessary or not appropriate[d]	Increase
	Physiotherapy referral made	Guideline	Routine EMR	Recorded referral to physiotherapy	Increase
Pharmacological interventions	Consideration of paracetamol use	Review	Template	Recorded as tried, offered, or declined full dose; or not appropriate[f]	Increase
	Paracetamol prescribed	Review	Routine EMR	Recorded prescription	Increase
	Consideration of topical NSAID use	Guideline	Template	Recorded as tried, offered or declined full dose; or not appropriate[f]	Increase
	Topical NSAID prescribed	Guideline	Routine EMR	Recorded prescription	Increase
	Oral NSAID prescribed	Guideline	Routine EMR	Recorded prescription	Decrease
	Oral NSAID prescribed in the presence of a relative comorbid contraindication	BNF	Routine EMR	Recorded prescription	Decrease
	Gastroprotection (PPI use with oral NSAIDs)	Review	Routine EMR	Recorded prescription (if oral NSAID prescribed)	Increase
	Opioid prescribed	Guideline	Routine EMR	Recorded prescription	Decrease

[a]Systematic review,[107] NICE guideline[5], or British National Formulary (BNF);[217] indicators taken from routine record had to be within 14 days of a clinical OA consultation; [b]compared to control group; [c]none, mild, moderate, severe; [d]not this time or no entry indicates non-achievement; [e]in those with last recorded BMI ≥ 25kgm^{-2} in up to previous three years; [f]unknown or no entry indicates non-achievement.

Due to a lack of established longitudinal monitoring of routinely-recorded care for OA, population benchmarks for the quality indicators cannot be set.

As previously stated, people with no record of a quality indicator outcome were assumed not to have that outcome, both for the template and routinely recorded management. The weight advice outcome assessment was restricted to patients known to be overweight at the index consultation. Oral NSAID use in the presence of relative contraindications was restricted to patients with a relevant recorded comorbidity in the previous two years. Gastroprotection was restricted to those patients with a record of a relevant oral NSAID prescription.

Sample size

The sample size required for the trial was based upon the requirements of the clinical effectiveness analysis, determined by the trial statisticians and reported elsewhere (Dziedzic et al., in submission). On the basis of a projected 10% annual consultation rate for clinical OA in patients aged 45 years and over,[21,311] the anticipated total population for this age group in eight practices (30,000 people) was expected to yield 3,000 OA consultations per annum. Accounting for repeat consultations, 2,000 unique patients were expected to consult in a six-month period, equating to approximately 1,000 patients per study arm. Allowing for an expected 50 clinicians consulted per arm, and an intraclass correlation coefficient (ICC) of 0.1, it was estimated that an OR of 1.6 in the dichotomous outcomes could be identified with 80% power and significance level (α)=0.05. The projected ICC used was similar to that reported for recorded items of advice as identified by Russell et al.[312] but rather larger than the majority of those reported by Adams et al.[313] for diverse variables in 31 primary care studies.

Blinding

The candidate and the trial statisticians were kept blind to the intervention or control group allocation until after the analysis was complete.

The index consultation

As in previous analyses, patients were allocated to an index consultation, namely the first consultation for any clinical OA in the trial period (codes starting N05.. formal OA diagnosis, or selected joint pain codes as in Chapter Three, section 3.4.3). The clinician consulted at this consultation was the index clinician. The index consultation was used to determine the start point for outcome success (120 days). In addition to the index clinician, BMI status and total morbidity were determined as the closest prior record and number of BNF subchapters prescribed in the 12 months prior,[218] respectively.

8.2.2 Statistical analysis

Baseline characteristics of each trial arm are first described in terms of demographics and the patient-level characteristics used as independent variables in previous chapters.

Baseline template use by practice within each trial arm was determined. This was used to facilitate a comparison between intervention and control arms, given the known variation in use of the template described in Chapter Six.

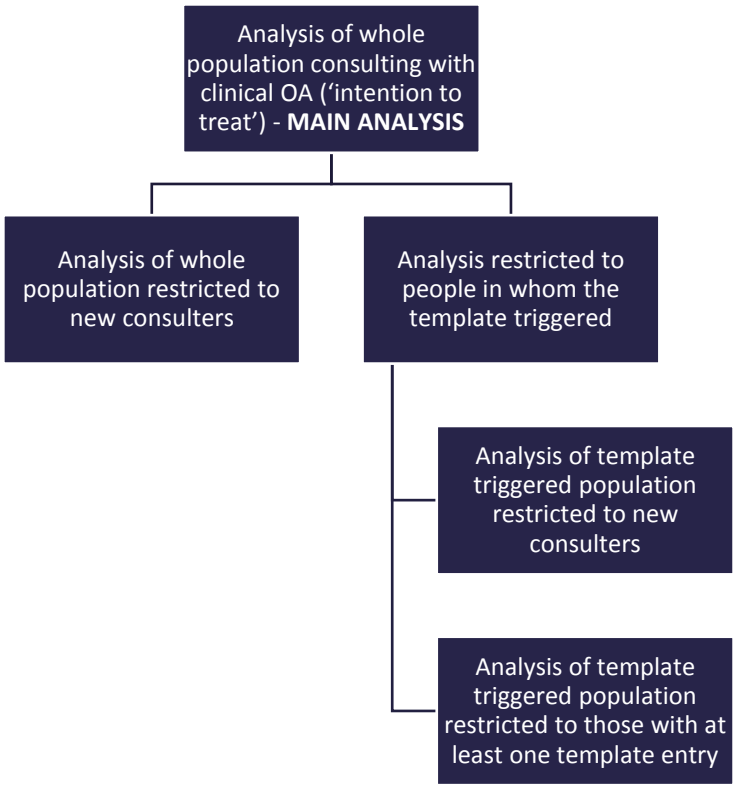
Bias is a recognised hazard in cluster trials.[314] The interpretation of cluster trials can be affected by an inability to include non-recruited participants in the analysis.[163] This analysis, however, was analogous to an intention to treat analysis by including all patients aged 45 years and over who consulted for clinical OA during the trial period, thereby substantially reducing the risk of selection bias. Intention-to-treat analysis in this context meant that patients were analysed as part of their practice randomised arm regardless of whether the individual patient received that arm's allocated care and regardless of whether template successfully fired. This approach was chosen as the aim of the study was to assess how quality indicators could be used (in a cluster trial) to assess the effect of an intervention. Just as with an intention to treat analysis in an individual patient-level randomised controlled trial, it was important to assess the outcomes on the basis of the way in which intervention and control arms had been assigned, not correcting post hoc for any problems in the way in which the indicators and recording template had been implemented.

Just as the in baseline period, the template did not trigger for every eligible person (see Chapter Six, section 6.4.1); non-triggering would have made it unlikely that the codes contained within the template would have been recorded as the codes were non-standard and only easily recorded via the template itself.

A main analysis of the trial period outcomes and several sensitivity analyses were undertaken, as shown in Figure 8-1. The first part of the main analysis examined overall template use. Associations of the intervention with recording at least one entry, achievement of at least one template indicator, and achievement of all template indicators were determined. The analysis used a two-level (patients within clinicians) multilevel logistic regression model. Adjustments were made for diagnostic category (OA or joint pain), sex, age band, and variables found to be significantly associated with quality achievement in Chapter Six (site of disease, BMI status, morbidity as measured by the BNF chapter count, multiple clinical OA consultations, and staff index consultation count) and baseline achievement of outcomes at a practice level. As baseline practice achievement was included in the model, practice was not separately included as an independent variable (although inclusion of practice as an additional independent variable was tested, it gave rise to model convergence problems and so was excluded from the final model). The second part of the main analysis explored associations of the intervention with the individual quality indicators. Multilevel logistic regression analysis was again used to compare the intervention and control arms on the quality indicators described, as dichotomous outcomes, adjusting for the same covariates as the first analysis.

As well as the intention-to-treat analysis, sensitivity analyses were undertaken in (i) those patients in whom the template was triggered in at least one consultation for clinical OA in the trial period, (ii) new consulters (defined as first clinical OA consultation since the introduction of the template and with at least 365 days since any previous OA or joint pain consultation) with at least one template entry, and (iii) those in whom at least one entry was made in the template in the trial period.

Figure 8-1: Main analysis and subpopulation analyses outline



The main analyses reported were two-level multilevel logistic regression models (patients within clinicians, though three-level models (patients within clinicians within practices) were also undertaken as sensitivity analyses. Analyses were undertaken through MLwiN 2.34[234] via the runmlwin[235] command in Stata 14.1.[298] Where possible, second-order penalised quasi-likelihood (PQL2) approximations were used; in some models, this was not possible due to model convergence problems, and other specified approximations were used instead. Results are presented as ORs with 95% CI. Results are reported to three significant figures.

8.3 Results

8.3.1 Trial population characteristics

The characteristics of the intervention and control arm aggregated practice populations consulting for clinical OA in the trial period are shown in Table 8-2. As expected (due to differences in practice population sizes) there were more clinical OA consulters in the intervention compared to the control arm (1110 to 836). There was substantial variation in the proportions of people with an OA

diagnostic code between arms (44.4% to 28.2%). There was some, but less pronounced, variation between arms in other characteristics including BMI status, and morbidity burden.

Table 8-2: Characteristics of clinical OA consulters aged 45 years and over, trial period

		Intervention	Control
Total		1110	836
Diagnostic group, n (%)	OA	493 (44.4)	236 (28.2)
Sex, n (%)	Female	652 (58.7)	511 (61.1)
Age band, n (%)	45-64	531 (47.8)	370 (44.3)
	65-74	268 (24.1)	232 (27.8)
	75-84	220 (19.8)	180 (21.5)
	85+	91 (8.2)	54 (6.5)
Site, n (%)	Knee	548 (49.4)	393 (47.0)
	Hip	219 (19.7)	184 (22.0)
	Ankle/foot	62 (5.6)	73 (8.7)
	Wrist/hand	104 (9.4)	103 (12.3)
	Unspecified	66 (5.9)	45 (5.4)
	Multiple	111 (10.0)	38 (4.5)
BMI status, n (%)	Not overweight	221 (19.9)	146 (17.5)
	Overweight	340 (30.6)	256 (30.6)
	Obese	383 (34.5)	246 (29.4)
	Unknown	166 (15.0)	188 (22.5)
BNF chapter count group, n (%)	0-4	462 (41.6)	425 (50.8)
	5-9	353 (31.8)	236 (28.2)
	10+	295 (26.6)	175 (20.9)
Number of clinical OA consultations, mean (SD)		1.53 (0.94)	1.79 (1.24)
New consulters, n (%)		574 (68.7)	741 (66.8)
Multiple consulter with clinical OA, period 4, n (%)		295 (35.3)	465 (41.9)
Recorded X-ray in period 4[a], n (%)		10 (7.0)	48 (12.4)
Number of index consultations undertaken by the consulted clinician, n (%)	Above the median	988 (89.0)	763 (91.3)
Patients with a comorbidity as a relative contraindication to oral NSAIDs in previous 2 years		431 (38.8)	278 (33.3)

[a]Relevant recorded X-ray within trial period 4, before the index consultation

8.3.2 Recorded achievement by trial arm, baseline period

As shown in Table 8-3, there was substantial variability between trial arms in the levels of outcome achieved in the six-month baseline period with generally higher levels in the intervention arm. There were uniformly low levels of written information and advice recorded. Recorded prescription outcomes were more similar between arms, except for recorded PPI use in people prescribed an oral NSAID (13.0% to 1.7%).

Table 8-3: Outcome achievement by trial arm in trial baseline period

Outcome measure	Number of patients with outcome in period 3, pre-randomisation *n* (%)	
	Intervention arm	**Control arm**
Total	977	747
Assessment		
Pain assessment	704 (72.1)	388 (51.9)
Function assessment	688 (70.4)	382 (51.1)
Weight record	275 (28.1)	149 (19.9)
X-ray recorded	242 (24.8)	21 (2.8)
Core interventions		
OA information provision	575 (58.9)	274 (36.7)
Written OA information	36 (3.7)	6 (0.8)
Exercise advice provision	579 (59.3)	285 (38.2)
Written exercise advice	38 (3.9)	8 (1.1)
Weight loss advice provision[a]	311 (52.9)	158 (35.1)
Written weight loss advice[a]	7 (1.2)	1 (0.2)
Non-pharmacological management		
Consideration of physiotherapy referral	425 (43.5)	192 (25.7)
Physiotherapy referral made	90 (9.2)	34 (4.6)
Pharmacological management		
Consideration of paracetamol use	622 (63.7)	348 (46.6)
Paracetamol prescribed	154 (15.8)	145 (19.4)
Consideration of topical NSAID use	537 (55.0)	295 (39.5)
Topical NSAID prescribed	255 (26.1)	174 (23.3)
Oral NSAID prescribed	176 (18.0)	108 (14.5)
Oral NSAID prescribed in the presence of a relative comorbid contraindication	39 (11.8)	73 (15.6)
Gastroprotection (PPI use with oral NSAIDs)	3 (1.7)	14 (13.0)
Opioid prescribed	331 (33.9)	226 (30.3)

[a]in those known to be overweight at the index consultation

8.3.3 General patterns of template use, trial period

As shown in Table 8-4, there was wide variation between practices and between arms in the frequency of template use and achievement of quality indicators derived from the template during the trial period. Between 38% and 80% of patients had at least one template indicator achieved across the eight practices, and this was reflected in a substantial difference between control (39.2%) and intervention (56.8%) arms. There was also a wide variation in the proportions of patients with all eligible template indicators achieved (<5% to >40% between practices).

Although there was a greater level of template use in the intervention arm during the trial period, the adjusted ORs (which adjusted for baseline achievement) for patients having at least one entry

Table 8-4: Recorded template use and proportions of template-derived quality indicators achieved in intervention and control practices, trial period

		Practice ID										
		Intervention					Control					
		1	2	7	8	Total	3	4	5	6	Total	Adjusted OR[a] (95% CI)
All patients	% patients for whom template fired	95.1	99.2	89.3	98.5	94.9	82.8	90.4	93.5	97.8	89.9	-
	% patients with at least 1 template entry	50.1	79.5	48.5	81.5	56.8	26.9	37.9	38.4	77.2	39.2	1.47 (0.61,3.54)
	% patients with at least 1 template indicator achieved	50.1	79.5	48.5	81.5	56.8	26.9	37.9	38.4	77.2	39.2	1.47 (0.61,3.55)
	% patients with all eligible template indicators achieved	23.1	12.6	10.2	40.0	21.5	4.6	19.3	7.3	37.0	13.8	1.58 (0.65,3.85)
Restricted to patients in whom the template fired	% patients with at least 1 template entry	52.6	80.2	54.3	82.8	59.8	32.5	41.9	41.0	78.9	43.6	1.48 (0.68,3.65)
	% patients with at least 1 template indicator achieved	52.6	80.2	54.3	82.8	59.8	32.5	41.9	41.0	78.9	43.6	1.49 (0.61,3.66)
	% patients with all eligible template indicators achieved	24.2	12.7	11.4	40.6	22.6	5.6	21.3	7.8	37.8	15.3	1.59 (0.65,3.90)

[a]Reference category is the control arm; adjusted for all covariates.

and recorded as having all template indicators achieved showed a trend of a positive effect of the intervention on quality of care but the results were not statistically significant. This remained true when the analysis was restricted to patients for whom the template fired.

23.8% of clinicians in the intervention arm made no entry in the template for any of their patients; 15.9% made at least one entry (and achieved at least one template-derived quality indicator) on every patient. For the control arm, the corresponding figures were 34% and 4%. In the intervention arm, 49.2% of clinicians did not achieve all template indicators in any patient, and 1.6% achieved all in every patient. For the control arm, 58% of clinicians did not achieve all template indicators in every patient and 2% did in all patients.

8.3.4 Intention-to-treat analysis of individual quality indicators

The outcome percentages for the individual indicators in each arm, with OR unadjusted and adjusted for all covariates can be seen in Table 8-5. Although the unadjusted odds for many outcomes were statistically significant, once adjustment for the relevant covariates was made, many of these became non-significant.

Compared to the levels of outcome achieved in the baseline period three (Table 8-3), population levels of recorded quality achievement generally fell in the trial period. There was generally a higher level of baseline achievement in the intervention practices and this persisted into the trial period. The recorded levels of the core interventions of information provision, exercise advice, and weight loss advice for those known to be overweight were similar in both arms with a non-significant trend toward greater use in the intervention arm. However, when restricted to written provision, written information (adjusted OR 23.2, 95% CI 7.10,75.8) exercise advice (adjusted OR 21.4, 95% CI 6.69,68.8) and weight loss advice in people known to be overweight (adjusted OR 23.5, 9% CI 4.61,120, PQL1 model) all strongly favoured the intervention arm.

Use of written information and advice was relatively stable in the control arm at 0.2% to 1.1% at baseline and 0.4% to 1.4% in the trial period; however recorded use of the various written materials increased substantially in the intervention arm (from 1.2%-3.9% to 13.7%-26.3%). Use of gastro-

Table 8-5: Adjusted odds of specified outcomes in intervention practices compared to control, trial period (95% CI). Intention-to-treat analysis.

Outcome measure	Number of patients with outcome *n* (%)		OR (intervention vs. control arm) (95% CI)		
	Intervention arm	Control arm	Unadjusted two-level model[d]	Adjusted two-level model[d]	Adjusted two-level model[d] in new consulters[e] only (*n*=1323)
Total	1110	836			
Assessment					
Pain assessment	612 (55.1)	317 (37.9)	2.69 (1.39,5.20)	1.34 (0.54,3.32)	1.60 (0.54,4.68)
Function assessment	606 (54.6)	307 (36.7)	2.71 (1.41,5.22)	1.14 (0.46,2.82)	0.69 (0.27,1.79)
Weight record	305 (27.5)	143 (17.1)	1.78 (1.07,2.97)	1.15 (0.66,2.03)	0.93 (0.53,1.64)
X-ray recorded	163 (14.7)	47 (5.6)	3.43 (1.68,7.00)	0.43 (0.09,1.99)	0.34 (0.08,1.52)
Core interventions					
OA information provision	549 (49.5)	267 (31.9)	2.97 (1.52,5.80)	1.31 (0.56,3.03)	0.78 (0.33,1.85)
Written OA information	292 (26.3)	12 (1.4)	25.7 (8.89,74.4)	**23.2 (7.10,75.8)**	**24.6 (6.80,89.4)**
Exercise advice provision	522 (47.0)	245 (29.3)	2.63 (1.45,4.77)	1.50 (0.68,3.29)	0.90 (0.42,1.92)
Written exercise advice	230 (20.7)	7 (0.8)	27.6 (8.57,89.1)	**21.4 (6.69,68.8)**	**11.7 (4.14,32.8)**
Weight loss advice provision[a]	327 (45.2)	130 (25.9)	3.08 (1.52,6.24)	1.39 (0.66,2.96)	0.98 (0.46,2.09)
Written weight loss advice[ab]	99 (13.7)	2 (0.4)	24.8 (5.47,113)	**23.5 (4.61,120)**	**12.9 (2.32,71.8)**
Non-pharmacological management					
Consideration of physiotherapy referral[b]	94 (8.5)	65 (7.8)	2.17 (1.18,3.97)	1.41 (0.56,3.57)	0.99 (0.39,2.48)
Physiotherapy referral made	109 (9.8)	19 (2.3)	7.66 (3.07,19.1)	**5.50 (2.13,14.2)**	**5.73 (1.74,18.9)**

(con't)

Outcome measure	Number of patients with outcome *n* (%)		OR (intervention vs. control arm) (95% CI)		
	Intervention arm	Control arm	Unadjusted two-level model[d]	Adjusted two-level model[d]	Adjusted two-level model[d] in new consulters[e] only (*n*=1323)
Pharmacological management					
Consideration of paracetamol use	549 (49.5)	282 (33.7)	2.69 (1.49,4.86)	1.42 (0.67,3.04)	0.97 (0.44,2.14)
Paracetamol prescribed	241 (21.7)	117 (14.0)	1.77 (1.32,2.39)	↑ **1.58 (1.14,2.20)**	↑ **1.70 (1.10,2.62)**
Consideration of topical NSAID use	496 (44.7)	274 (32.8)	2.00 (1.14,3.53)	0.91 (0.43,1.92)	0.56 (0.26,1.23)
Topical NSAID prescribed	326 (29.4)	184 (22.0)	1.46 (1.01,2.10)	1.13 (0.75,1.70)	1.03 (0.63,1.67)
Oral NSAID prescribed	176 (15.9)	136 (16.3)	1.26 (0.99,1.60)	0.79 (0.53,1.16)	1.05 (0.66,1.67)
Oral NSAID prescribed in the presence of a relative comorbid contraindication	52 (12.1)	43 (15.5)	0.74 (0.41,1.33)	0.74 (0.42,1.32)	0.88 (0.43,1.82)
Gastroprotection prescribed (PPI)[c]	69 (39.2)	49 (36.0)	0.90 (0.60,1.35)	0.96 (0.50,1.82)	1.33 (0.43,4.16)
Opioid prescribed	364 (32.8)	228 (27.3)	1.05 (0.62,1.78)	0.85 (0.66,1.09)	0.87 (0.64,1.17)

Adjusted for OA or joint pain code, sex, age band, site of disease, BMI status, morbidity load (BNF chapter count), multiple clinical OA consultation, staff member index consultation count dichotomy, and practice pre-trial achievement; [a]in those known to be overweight at time of index consultation; [b]PQL1 model; [c]denominator: those prescribed an oral NSAID; [d]patients within clinicians; [e]new consulters defined as first clinical OA consultation since the introduction of the template and with at least 365 days since any previous OA or joint pain consultation

protection increased in the control arm (1.7% up to 5.9%) but diminished in the intervention arm (dropping from 13.0% to 6.2%).

Achievement of assessment measures were more similar between intervention and control arms for pain and function, and weight recording (all of which were recorded in the e-template). X-ray use fell in the intervention arm (24.8% to 14.7%) and increased from 2.8% to 5.6% in the control arm with adjusted OR 0.43 (95% CI 0.09,1.99). However, statistical significance was not achieved.

Physiotherapy referral was recorded as considered (captured by the template) at a similar level in each arm, though there was a non-significant trend toward increased consideration in the intervention arm (adjusted OR 1.41, 95% CI 0.56,3.57). Recorded use of actual referral to physiotherapy was significantly greater in the intervention arm (adjusted OR 5.50, 95% CI 2.13,14.2).

There were greater odds of consideration of paracetamol via the template in the intervention arm compared to control, though this did not achieve statistical significance (adjusted OR 1.42, 95% CI 0.67,3.04), and greater odds of a recorded actual prescription (adjusted OR 1.58, 95% CI 1.14,2.20). There were reduced odds of prescription of oral NSAIDs in the intervention arm (adjusted OR 0.79. 95% CI 0.53,1.16) and for reduced prescription of oral NSAIDs amongst people with a recorded comorbidity that was considered to be a relative contraindication to their use (adjusted OR 0.74, 95% CI 0.42,1.32) but neither achieved statistical significance.

The analysis was repeated as a three-level model of patients within clinicians within practices. This did not demonstrate any statistically significant explanation of residual variance at the practice level in the adjusted model. The effects on the OR estimates showed a moderate increase for the written information (adjusted OR up to 29.5 from 23.2) and advice outcomes (for exercise, increasing from 21.4 to 25.5 and for weight loss in people known to be overweight from 23.5 to 28.8) but no change for physiotherapy referral or paracetamol prescription.

Restriction of the eligible population to new consulters only did not substantially change the results, with all significant associations remaining so.

8.3.5 Patients in whom the template triggered

Restriction of the analysis to those patients in whom the template was triggered (n=1804, 92.7%) did not change the findings. A further restriction to new episode consulters only was undertaken. This showed a tendency for the control arm to have generally slightly higher frequencies of the specified outcomes and for the intervention arm to have slightly lower frequencies. The achievement of statistical significance of the ORs was unchanged, though there was slight fluctuation in the estimates except for written weight loss advice, which sustained a big drop in the estimated adjusted OR.

8.3.6 Patients in whom at least one template entry was made

As shown in section 8.3.3 above, there was a marked discrepancy between intervention and control arms in the frequency with which at least one template entry was made, with the control arm achieving at least one entry in 39.2% of patients whilst the intervention arm achieved this in 56.8% of patients.

Pain and function assessments were the most frequently-completed items on the template, with completion rates (when at least one entry was made) of over 90% in both trial arms. This led to problems with the covariates chosen in the models with pain and function as the outcomes. The intended model had to be discarded in favour of a restricted model which included only diagnostic category, sex, age at index consultation (continuous variable), and practice baseline achievement. Other outcomes were assessed using the standard model.

Findings were similar to previous analyses except for written weight loss advice in people known to be overweight, which dropped to 17.6 from 34.3 in the analysis of template-triggering patients (Appendix H.2).

When restricted to those patients with at least one template entry (Appendix H.3) there was no statistically significant explained variance at the practice level. For written information, the adjusted OR estimate increased to 29.5 from 24.5 and for written exercise advice from 21.8 to 23.5.

8.4 Discussion

8.4.1 Main findings

The objective of this chapter was to describe how the effects of a complex intervention to improve uptake of NICE-recommended management strategies for OA in general practice registered populations could be assessed through quality indicators including those recorded via a template integrated to the clinical information system.

This analysis of the MOSAICS trial period data has shown that the model OA consultation intervention was associated with substantial improvements in some important components of the primary care of OA as recorded within the template: written information and written advice about exercise and weight loss, recorded physiotherapy referral, and paracetamol prescription.

However, the intervention arm had higher levels of template use (as measured both by template triggering and by recording of at least one item in the template) at baseline and during the trial period, though there was a fall in template use between the baseline and trial periods in both arms. Wide variation during the baseline period between clinicians, practices, and trial arms caused difficulty in estimation of the effects in the trial period. The effect of item order seen in the assessment of the template (in which the final template items were less frequently completed than the initial items) remained a problem with the biggest falls in recording in the domains listed later in the template. Given this, it is unclear to what extent the greater use of the template in the intervention arm in the trial period was due to the greater level of use in the baseline period or whether the MOAC intervention helped sustain template usage. The signal from the findings relating to the template use measures might be taken to imply some sustaining effect on template usage from the intervention, though statistical significance was not achieved and caution is needed in drawing conclusions.

Stratified block-randomisation was used to allocate practices to intervention or control arms. The randomisation process was conducted by computerised random number generation in accordance with good research practice by the Keele CTU, independently of the trial investigators. In common

with the trial statisticians, the candidate was kept blind to the allocations until after the outcomes analyses.

With any randomisation process, an imbalance of baseline characteristics can occur purely by chance and this is presumed to be the cause of the differences in baseline characteristics between the trial arms, with the small numbers of practices in each arm (four clusters per arm has been suggested to be the minimum number required[315,316]).

The sample size for the MOSAICS trial was calculated to meet the requirements of the main outcome (patient-level clinical effectiveness measures). On the basis of a projected 10% annual consultation rate for clinical OA in patients aged 45 years and over,[21,311] the anticipated total population for this age group in 8 practices (30,000 people) was expected to yield 3,000 OA consulters per annum. Accounting for repeat consultations, 2,000 unique patients were expected to consult in a six-month period, equating to approximately 1,000 patients per study arm. Allowing for an expected 50 clinicians consulted per arm, and an intraclass correlation coefficient (ICC) of 0.1, it was estimated that an odds ratio of 1.6 in the dichotomous outcomes between trial arms could be identified with 80% power at a significance level (α)=0.05. The projected ICC used was similar to that reported for recorded items of advice as identified by Russell et al.[312] but rather larger than the majority of those reported by Adams et al.[313] for diverse variables in 31 primary care studies.

Actual variability between clinicians was greater than expected for some indicators. For example, the odds ratio for recorded pain assessment in the trial period (in new consulters only) between intervention and control arms was estimated to be 1.6 but this did not achieve statistical significance (95% CI 0.54,4.68), suggesting that the study may have been under-powered for some indicators. However, this analysis was intended to test the feasibility of quality indicators as trial outcomes and in this context, the ability of the indicators to discriminate between trial arms on several aspects of care (significant differences in written information/advice, paracetamol prescription, and physiotherapy referral) suggests potential for indicators to be used in this way.

One conclusion from this analysis is that researchers should use conservative estimates of the ICC when undertaking sample size calculations for clustered designs which examine process measures, such as in this study.

The increased use of written information seen in the intervention arm was likely to represent the use of the OA guidebook provided to the intervention arm only (though other printed materials were available by request to the control arm for use had the practices so chosen).

There was a trend toward reduced use of X-ray investigations and, although this did not achieve statistical significance, the apparent direction of change was in accordance with the NICE guidance.[5] There was a small reduction in odds of an oral NSAID prescription in the intervention arm, which again would be consistent at a population level with the NICE guidelines to use other analgesics in preference to oral NSAIDs. There was also a trend toward lower oral NSAID prescription amongst people with a recorded comorbidity that presented a relative contraindication to such prescription. Again, this did not achieve statistical significance and it is not clear whether this was a chance finding, or whether the study lacked power to detect a real difference.

8.4.2 Comparison with previous literature

Despite a fall in the recorded levels of achievement, levels remained higher in the intervention arm than those reported in prior literature. For assessment of pain and function, as the most frequently-completed items on the template, rates of recorded quality indicator achievement of 55% in the intervention arm (37%-38% in the control arm) were lower than achieved in the baseline period[99] but still exceeded those reported by Broadbent et al.[69] (27% for pain and 43% for function). Rates of information provision at 50% (26% for written) in the intervention arm also exceeded the 30% reported by Broadbent et al.,[69] and was similar to the 55% (patient self-reported) found by Østerås et al.[317] though rates in the control arm at 32% (1.4% written) were more in line with Broadbent's assessment. Provision of exercise information at 47% (intervention) and 29% (control) compared well to the 25% reported by Li et al.,[80] and for the intervention arm at least were comparable to the 49% reported by Østerås et al.[82] Advice about weight loss was similar at 45% compared to 46%

(patient self-reported) in Østerås et al.[317] but lower than a prior estimate by the same author (68%[82]). Recorded physiotherapy referral at 31% compared favourably to the estimate of 25% of people advised to exercise (which included physiotherapy referral) by Li et al.[80]

In Chapter Seven, it was shown that the template was associated with an increased recorded prescription frequency for paracetamol and topical NSAIDs.[99] This prior association meant that it was additionally difficult for the MOSAICS intervention to demonstrate an effect in these domains. The recorded frequency of patients receiving a prescription for paracetamol within 14 days of a clinical OA consultation that occurred within 120 days of the index consultation (22% in the intervention arm and 14% in control) has no direct comparator in the prior literature: Broadbent et al.[69] examined recorded first use of paracetamol, identifying a rate of 48%, and Steel et al. 41%.[70] However, there was not the same time period restriction to those analyses.

8.4.3 Use of quality indicators as outcomes in clinical trials

Quality indicators have an established role in service evaluation and audit. However, unlike many established outcome measures, quality indicators do not have a well-established use in clinical trials for assessment of outcomes. Eyssen et al.[318] used indicators as secondary outcome measures in an assessment of occupational therapy in multiple sclerosis care. The CARAT trial for diabetes care used several diabetes quality of care indicators as outcome measures (nephropathy screening, retinopathy screening, foot examination, and peripheral neurological testing).[319] Again in the field of diabetes, Shah et al.[320] used diabetes quality indicators as outcomes in a cluster trial of an educational toolkit. Performance indicators have also been compared to outcome indicators in a study of major trauma incident exercises.[321] The protocol for a randomised controlled trial of a quality initiative in chronic obstructive pulmonary disease also states that quality indicators will be used as outcome measures.[322] The use of quality indicators as outcomes in this trial of a complex intervention to increase adherence to the NICE OA guidelines extends the knowledge base for use of indicators as trial outcomes. Quality indicators often refer to processes of care.[44,124] Where such

processes are considered intrinsically desirable, it should generally be feasible to use indicators as outcome measures in implementation trials.

However, there are issues which need to be overcome. In trials that seek to improve clinical outcomes in osteoarthritis, further work to understand how well recorded process measures are linked to clinically-meaningful outcomes would be required before indicators could be used in place of more traditional outcome measures.[46,107] This might be particularly true for core interventions such as education/information and lifestyle advice (exercise, weight loss) where the content of advice would be expected to vary across clinicians and potentially by patient characteristics.

Due to the e-template not firing in all patients (5% missing in the intervention arm and 10% in the control arm), some information is likely to have been lost since it seems probable that clinicians would have recorded at least some information on some of those patients in whom the e-template did not trigger. This loss of information may cause a bias in treatment effect if there is a differential rate of non-triggering between intervention and control arms as happened here. The consistent satisfactory functioning of the e-template is an important consideration for any future trials using indicators as their primary outcome measures.

The high baseline achievement, and baseline variability between trial arms and between clinicians may have reduced the power to detect a difference in some of the outcome measures. The use of the e-template was lower in this trial period compared to the baseline period. It is possible that this was the result of so-called 'initiative fatigue', such as has been reported for the QOF.[323-325]

Multiple comparisons were unavoidable when assessing multiple indicators of quality of care. In the MOSAICS study, given the consistency of the findings across the written information and advice indicators, and the strength of the physiotherapy referral association with the intervention, a genuine trial effect seems likely. The effect on paracetamol prescription was less dramatic but given the routinely-recorded nature of this outcome, its promotion by the NICE 2008 guidelines, and a prior increase associated with the template introduction, a genuine increase is again plausible.

8.4.4 Conclusion

In principle, quality indicators for OA management identified from a systematic review and supported by an enhanced recording method in general practice should be implementable as outcome measures in an intervention trial for OA management. However, the introduction of an indicator recording method itself (the template) was found to have some effects on recorded quality of care, and this, coupled with high levels of achievement at baseline in both trial arms for some processes of care (such as pain and function assessment), and baseline variation in indicator achievement limited the scope of the study to detect a difference between arms.

Quality indicators may function better as trial outcome measures in trials that have a greater number of randomly allocated clusters, such that the differences in baseline achievement between intervention and control arms may be less pronounced. Whilst indicators could in principle be used in trials that use individual-level randomisation, it is likely that relatively few interventions that are likely to have an effect on recorded processes of care would be suitable for individual patient-level randomisation.

Chapter Nine will summarise the findings across the thesis as a whole, and link these to the original aims and objectives. The overall conclusions arising from the work are set out, with proposals for implementation of the findings and implications for future research.

Chapter Nine: Discussion and conclusions

The overall aim of the work described in this thesis was to identify indicators of quality of care for OA in general practice, to assess how they night be implemented in a general practice EHR, and to test their feasibility as outcome measures in a clinical trial. This chapter will summarise the main findings from the research and link these to the initial aims and objectives. The strengths and limitations of the work overall will be discussed, and future implications and conclusions considered.

9.1 Summary of findings

9.1.1 Identification and synthesis of indicators

Objective (i) to review the existing evidence for quality indicators for OA, applicable to general practice

Quality indicators for the primary care of OA were identified from a systematic review, grouped into assessment and treatment themes derived from the NICE OA management guidelines, and synthesised into indicators considered suitable for implementation in UK general practice. Overall, fifteen OA indicators were considered suitable for implementation. A need for additional indicators relating to treatment (topical NSAIDs) and to certain 'do not do' indicators (such as referral for arthroscopic lavage) was identified.

9.1.2 Development of a template for OA consultations

Objective (ii) to identify how the quality indicators may be assessed through general practice medical records and develop a mechanism for enhanced recording where necessary

The need for additional data capture in general practice to enable routine measurement of the quality indicators was identified, concluding that only measures of prescribing and referral could be used without improved data recording.

An enhanced recording tool, the OA template, was developed and implemented. This single-screen (10-item) template was produced to facilitate recording of the main items of care for OA that were

not well-covered by the routine record: assessment of pain and function, a record of or advice about relatively safe pharmacological options (paracetamol, topical NSAIDs), information provision and advice about exercise and weight loss (where relevant) and consideration of physiotherapy referral.

A variable response rate to the template questions was identified in the initial six-month use of the template (Chapter Six), possibly related to the question order in the template may have affected response rates. There was considerable variation between clinicians in whether or not the template was used, and if it was used, how fully-completed it was. The high levels of completion achieved by some clinicians suggests that the template was feasible in principle but that additional work to understand how the variation could be minimised is needed.

9.1.3 The denominator population for quality indicators: a Read-code definition of clinical OA

The interpretation of quality of care depends upon the patient group to whom quality of care indicators are applied. A definition of clinical OA for the MOSAICS study was developed through a robust process involving an initial draft list of codes selected by the candidate which was then revised through a process of consensus, assessment in the locally-derived CiPCA database, and then finally a review by the MOSAICS research team. The process sought to strike a balance between practicality for clinicians and the informatics team (avoiding making the definition too sensitive) and on the other hand reducing the risk of selection bias in the MOSAICS study (through an overly specific set of codes that could have been avoided in clinical use). Residual concern about a lack of specificity in the clinical definition of OA used to trigger the template and assess study outcomes was considered further in the assessment of coding in the baseline phase one.

9.1.4 Naturally-occurring coding patterns for clinical OA

Objective (iii) to describe the naturally-occurring patterns of OA and peripheral joint pain morbidity coding in general practice

In the eight MOSAICS study practices, before inception of the study (phase one), the age-standardised diagnosed OA consultation prevalence over the 12-month period was 3.2%.

Factors associated with a diagnosis of OA rather than joint pain were found to be older age, unspecified or multisite disease, obesity, multiple clinical OA consultation, and X-ray use prior to the index consultation. Ankle/foot site of disease was associated with reduced odds of an OA diagnosis compared to joint pain compared to the knee. A hypothesis that patients with a higher total level of morbidity as assessed by BNF chapter count (as a measure of biological rather than chronological age) would be more likely to receive a diagnosis of OA was not found to be supported.

Investigation of diagnostic misclassification (attribution of a clinical OA label to people with joint pain who actually have non-OA joint pain), suggested that the true annual consultation prevalence rate for clinical OA lay between 4.7% and 7.2%, with a point estimate of 5.7%. Compared to the knee, ankle/foot and wrist/hand pain codes were least likely to represent clinical OA.

The restriction of the definition of clinical OA in this thesis to peripheral joints, excluding the axial skeleton, shoulder, and elbow, is considered likely to have increased the specificity of the definition of true OA. As noted in Chapter Four, additional improvement in diagnostic specificity may have been achieved by exclusion of the ankle/foot joints. It is unknown to what extent the recorded quality of care, and the effects of the template and model OA consultation, would have been altered by a change in the definition of OA. Given the identified link between an OA diagnosis and some aspects of recorded quality of care, it seems plausible that a clinical OA definition that included the axial skeleton, shoulder and elbow (with use of pain as well as formal OA Read codes) may have reduced the apparent effects of the template and trial intervention in the combined clinical OA population due to lower specificity for joint pain as a marker of OA at these sites.

9.1.5 Naturally-occurring coded quality of care for OA

Objective (iv) to describe the routinely-recorded and coded quality of care for clinical OA in general practice

Assessment of routinely recorded quality of care for OA in the eight MOSAICS practices during phase one (Chapter Five), before study commencement, identified very low levels of coded information relating to pain and function assessment and the core interventions of education and

advice regarding exercise and weight loss (where relevant). It was possible to measure quality indicators related to prescribing and referral activity.

Factors associated with better routinely-recorded quality of care were, for prescribing measures, those associated with increased risk of OA (female sex, older age), higher total morbidity, and multiple clinical OA consultation. Primary care referral was more likely for people with joint pain, whereas secondary care was more likely for people with diagnosed OA. The oldest patients had reduced odds of secondary care referral. Site had differing associations but generally patients with the ankle/foot, wrist/hand and unspecified sites had lower odds of referral. Patients with a recorded X-ray and multiple clinical OA consultations had greater odds of referral.

Increased age was associated with reduced odds of some assessment measures but increased prescribing processes and reduces referral processes, raising a question about equitable access to appropriate treatments for the very elderly.

9.1.6 OA quality of care measured through the template

Objective (v) to describe the coded quality of care for OA as captured by the enhanced recording mechanism

Pain and function assessments were relatively well-recorded in the initial six months after introduction of the template, at 63% and 62% of patients respectively. The rest of the indicators were achieved in the range of 45 – 56% with the exception of physiotherapy referral and recording of weight measurements. Achievement of the template-derived indicators was more common than the complementary routinely-recorded indicators (i.e. consideration of processes was more common than actual prescription or referral).

An OA diagnosis rather than joint pain was associated with achievement of all template indicators. Older age, multiple clinical OA consultation, and consultation with a clinician with a higher clinical OA workload (as measured by the index consultation count) were also associated with greater odds of achievement. Ankle/foot disease was associated with reduced odds of indicator achievement as were, though with less consistency, an unknown BMI status and higher total morbidity burden.

Prior recorded X-ray use was not associated with template-derived quality of care measures but a record of X-ray within 14 days was associated with various template measures.

9.1.7 Effects of the template on recorded quality of care

Objective (vi) to estimate the effect of the enhanced recording mechanism on quality of care determined by routinely-recorded and coded means

Amongst measures recorded through routine mechanisms in the EHR, the template was found to be associated with a rise in weight recording and a rise in use of the relatively safe pharmacological options (paracetamol, topical NSAIDs).

9.1.8 The implementation of quality indicators as outcomes in a clinical trial

Objective (vii) to investigate the feasibility of use of quality indicators in assessment of a cluster-randomised controlled clinical trial of a complex intervention to improve care for OA

The use of quality indicators as, and templates as a source of, outcomes in a cluster-randomised controlled trial of a model OA consultation intervention were investigated. Due to variation in baseline performance between the control and intervention arms in aspects of care measured by the e-template, assessment of change relating to the intervention itself was difficult. However, there was clear evidence of improved recording of written information and written advice to exercise and lose weight (where relevant), and evidence of increased physiotherapy referral.

9.2 Strengths and limitations

There were a number of important strengths to this work. It was founded on a rigorous systematic review and narrative synthesis which has since been cited by the NICE 2014 OA guidelines. The introduction of the indicators into the enhanced recording template was considered in the light of the best available evidence.

This was a large study, with 1946 patients included in the trial analysis stage, compared to 525 in the clinical effectiveness assessment (not reported here). Estimates of the effects of the template and the model consultation have been adjusted for a wide range of potential confounding factors,

especially through the use of a simple measure of total morbidity load (the BNF chapter count) and a measure of clinical OA workload for clinicians.

The data collection from the MOSAICS practices was designed to have a high sensitivity to clinical OA and thus avoid recording bias (by shifting diagnostic recording patterns) once the template had been introduced as well as during the cluster trial. A novel aspect of this work was the additional estimation of diagnostic misclassification through the extensive analysis of joint pain consultation narrative, giving rise to new estimates of the consultation prevalence of clinical OA, at approximately 53% to 80% of the consultation prevalence before misclassification assessment (point estimate of 64% of the original estimate).

In datasets that are time-limited (i.e. do not include the whole primary care record since first presentation of clinical OA), caution should be exercised about inference of individual-level quality of care. Routine determination of care processes such as exercise advice or referral at least once since diagnosis, or first-use of relatively safe analgesics options, cannot be made without consideration of the full record. The timeframe used for linking clinical activity to quality indicators was arbitrary, though in keeping with previous research practice.[264] There was a risk of misclassification in attribution of patients to prescribing and referral outcomes, as some patient not identified as having received such a process may actually have received one but outside of the 14 days after a consultation for clinical OA, for example, those on repeat prescriptions. Some patients receiving such processes within 14 days may in fact have received it for an unrelated condition. The direction and effect of such misclassification is not possible to determine with any certainty. There is no reason to suppose that it would have resulted in a bias in assessment of the trial outcomes, though the absolute levels of achievement of the prescribing indicators should be interpreted with caution. Referrals were more likely to be recorded within a relatively short space of time after a consultation as no similar repeat facility exists for referrals, and so it is considered that less misclassification results from the use of a 14-day period for referrals.

It would have been desirable to include additional variables in the models. For example, an OA severity measure would be relevant to adjust for the case-mix issue described in Chapter One. A health literacy measure would have been desirable to include to identify to what extent patients' health literacy influenced provision of care for OA as it has previously been found that there is an inverse relationship.[326] The risk of lack of health literacy increases with worse deprivation[327] but Broadbent[69] did not demonstrate a link between individual or practice-level deprivation and achievement of quality indicators for OA. Clinician factors such as those determined by Clarson et al.[328,329] to predict agreement about a QOF domain for OA (special interest in musculoskeletal disease, a higher research degree, familiarity with the NICE guidance) and OA monitoring (belief in the importance of monitoring, special interest in musculoskeletal disease, a Master's degree, familiarity with the NICE guidance, and working full-time) would also have been of potential importance in this study.

The MOSAICS study practices, drawn from a mix of urban, suburban, and rural areas of the North-West Midlands and Cheshire, were considered to form a representative sample of general practice. The generalisability of the findings may be affected by various factors, however. There was a substantial difference in the way in which practices coded clinical OA (as formally-diagnosed OA, or as peripheral joint pain), as described in Chapter Four. Practice 1, as the largest contributor practice, recorded more OA as a proportion of all clinical OA than the other practices. The study practices were all willing to participate in the research and received additional resources to participate in the study. It is uncertain to what extent the findings might be replicated in other general practices. An evaluation of a similar approach in routine practice forms part of the JIGSAW implementation, discussed in 9.4, Proposals for implementation.

The apparent selection bias potentially arising due to the differences in baseline characteristics (generally higher levels of achievement of core OA management processes in the intervention arm but a greater proportion of people with an OA rather than joint pain diagnosis in the control arm) should have been largely adjusted for through inclusion of baseline levels of achievement in the multilevel logistic regression models, as well as the inclusion of a diagnostic group (OA or joint pain)

variable. There remained a risk of unbalanced and unmeasured confounding factors between trial arms which may have affected the trial outcomes. The magnitude and direction of these effects are uncertain but are considered likely to have favoured the recorded quality of care in the intervention arm. Even within arms, there was substantial variation in the recorded proportions of people with OA or joint pain diagnoses by practice. Variation between practices is likely to be a typical finding. The variation seen between practices in this relatively small number of clusters does limit the generalisability of the findings.

The estimates of effects described were shown to be greater in patients for whom at least one template entry was recorded. This may reflect greater use of the template by clinicians for patients they considered to have clinically significant OA and as such, the analyses that restricted the eligible population to those with at least one entry may be a truer marker of the efficacy of the template and the model consultation.

As noted in Chapter Three, the use of a medical record review for assessment of quality of care has limitations predominantly relating to completeness of information and retrievability through use of codes.[171] It is uncertain to what extent recording reflects actual delivery of quality of care, as the analysis was dependent upon the coding of information from consultations and it is known that such information capture is not entirely reliable.[330] It has been identified that the content of the medical record may be at variance with the findings from standardised or observed consultations.[331-333] However patient-reported quality indicator achievement in MOSAICS[317] showed broadly comparable achievement rates for those indicators measured through both patient self-report and medical records.

9.3 Discussion

9.3.1 Quality indicators for the primary care of OA

A set of quality robust indicators for primary care were identified but implementation of these in routine practice without a change to recording practice does not seem to be possible. Through use of an enhanced recording template, there was substantially better coverage of the indicator set,

though recorded performance in the MOSAICS practices in the first six months of template use was variable between clinicians and practices. There was evidence of a reduction in template use in both arms during the trial period (second six-month period of use).

A lack of comprehensiveness of indicators (omissions related to the absence of indicators on topical NSAIDs and other NICE-recommended OA treatments, as well as to the absence of 'do not do' indicators), was perhaps in part related to a lack of continuous refinement of indicator sets in light of emerging evidence. The NICE quality standards[334] have partly addressed this but there remains a gap in coverage between the indicators (or standards) and recommended practice. Routinely collected information that could be used to assess indicators of OA care in Maxwell's domains of effectiveness, efficiency, equity, acceptability, access, and relevance[56] would be useful development for primary care. The indicators identified herein go some way toward measuring efficiency (as one indicator would measure how often all of the other indicators have been met prior to specialist referral). New indicators, including patient-reported process (and possibly outcome) measures would be needed to address the remaining domains for OA in primary care.

Limitations in the understanding of quality indicators themselves were identified, notably a lack of information about quality improvement resulting from implementation of such indicators (external validity). Some evidence suggests that in fact simple monitoring of quality indicators implemented in laboratory medicine did not consistently result in quality improvement.[335]

It is known that the medical record provides at best a partial picture of the consultation with inherent, systematic bias in coded information.[330 332] Marshall et al.[144] argued that some aspects of care were sufficiently important that a failure to capture them in the medical record would in itself be a marker of poor quality of care. However, the threshold for a minimally acceptable record is hard to define. Certainly, the evidence-based management of OA in general practice would add more time to a consultation compared to treating it as a low clinical priority. Levene et al. have proposed four principles for general practice data collection: to understand population health needs better, collection of only necessary data, rational use of data to help improve outcomes and

policies, and maximisation of practice engagement.[336] The relatively modest proposals for recording information in an OA consultation do not seem a disproportionate response to the scale of OA as a clinical problem.

The population approach to determination of quality of care as used in this work was in keeping with that adopted by the QOF. This uses process and surrogate or intermediate outcome indicators, measured at the level of the patient, to make payments to general practice on the basis of aggregated level of care achieved in the registered practice population and is used as a means of assessment of general practice, such as by the Care Quality Commission.[337] It was also consistent with that suggested by Gribben et al. for use in New Zealand primary care, which included population-level indicators such as ratios of lower risk drugs to higher risk.[338]

Overall, the indicators seem to be feasible for use but there is a need to understand how to make clinical engagement with their use more consistent between clinicians as well as within clinicians over time.

9.3.2 OA coding as an aspect of quality of care

One aspect of the medical-technical quality of care is the making of a correct diagnosis,[40,41,339] and the recording of adequate information in the health record.[340] By extension, one might also argue that the recording of a correct diagnosis is also an aspect of technical quality of care and this would be consistent with the first NICE OA quality standard about making a diagnosis of OA.[334] The potential benefits to the patient lie in provision of adequate information about the condition – which requires a working diagnosis to be made – and for sufficient continuity of information between clinicians.[341] There would also be potential benefits to clinicians (information continuity, again) and to public health and health care commissioning (for planning and service evaluation purposes, consistent with the need identified by Brand et al.[342]). Clinical audit depends on adequate information about the proportion of eligible people receiving care or their outcomes, which again requires sufficiently precise recording of diagnoses to enable this. This would be comparable to

QOF[102] for long term conditions such as diabetes, the care of which has been found to benefit from a structured approach as described by Brennan.[343]

There was no evidence that the template changed practice regarding the recording of a diagnosis of OA or joint pain. The diagnostic misclassification work has indicated that there is a need to improve coding of clinical OA. Restriction of the definition of OA to a formal diagnosis is too specific, with cases missed. On the other hand, the joint pain diagnoses used were too sensitive, with misclassification again, resulting in non-OA cases being included. A refined definition of clinical OA from diagnostic codes is needed, potentially excluding joint pain codes relating to the ankle/foot, wrist/hand and unspecified sites. Coding training and feedback, such as used in the inception of CiPCA,[22] may assist with diagnostic coding.

9.3.3 Quality improvement in OA

The association of prescribing patterns and some known risk factors for OA suggests that clinicians may have been more willing to treat people with clinical OA people pharmacologically if they fit the stereotype of a patient with OA. If so, increased education might be expected to help with management of clinical OA in a wider sense. Association between increased odds of many prescriptions and older age, coupled with reduced odds of referral to secondary care, may represent an appropriate desire at the level of either the clinician or patient to manage symptoms with analgesia rather than surgery, but it is also possible that this could have reflected inappropriate reduced access to effective secondary care treatments on ground of age.

The association of odds of referral with prior X-ray use suggests that clinicians selected patients for referral on the basis of radiographic reports: this may represent a desire for clinical certainty prior to referral, symptom severity as a confounding factor, or an attempt to ration care. Given the association of template-derived indicator achievement with X-ray use within 14 days after a clinical OA consultation, it is possible that, for some patients, clinicians undertook a more in-depth approach to assessment and management of clinical OA, including the use of radiographic assessment even though this is not a recommended investigation. If the triggers for this more

thorough approach could be identified, it is possible that the higher level of quality might be achievable even with reduced used of X-rays, through appropriate clinician education and development.

Had the intention been to implement the template as a quality improvement tool for OA in its own right, additional strategies could usefully have been considered such as additional training, customisation, and educational outreach. This approach would be more consistent with that adopted in the JIGAW implementations of the MOSAICS work, discussed under section 9.4, Proposals for implementation. Nevertheless, the template had a small-to-moderate effect on prescribing behaviours for the relatively safe pharmacological options. The model OA consultation was associated with increased use of written information and advice as well as physiotherapy referral and it is plausible that a template that instead acted more formally as a prompt to refer people might be capable of a similar effect.

The reason for a lack of identified change in other aspects of care for OA after the model OA consultation training may have been due to various factors including baseline variation in recorded quality of care and high levels of baseline achievement in some indicators. The *Good Life with osteoArthritis in Denmark* (GLA:D) programme, which has some similarities with the MOSAICS model consultation, is reported[344] [abstract] to have resulted in reduced levels of paracetamol, NSAIDs, or opioids after 3 months. A stepped care programme in the Netherlands (Beating osteoARThritis [BART][345]) demonstrated high levels of use of most of the first-tier interventions, other than glucosamine, after two years of the programme (education 82%, lifestyle advice 73%, paracetamol 83%); exercise (a second tier intervention) was used by 63% of patients after two years.[346] No control group was reported for either of these studies. It is not clear whether such success might be transferrable to a UK context but it would be consistent with the possibility that the lack of change in other indicators in the MOSAICS trial was due to other factors beyond lack of benefit. However, a previous UK general practice-based trial of a self-management programme for OA (involving a guidebook and up to six sessions of advice on self-management)[347] did not show any effect on pain, function or contact with primary care, though there was an improvement in

anxiety and self-efficacy. There remains a need to identify how to translate the best evidence on treatments for OA into routine clinical practice.

Potential targets for quality improvement initiatives are indicated by increased odds of achievement of relevant indicators with an OA diagnosis, multiple clinical OA consultation, and an index consultation with a clinician who conducted more than the median number of index consultations. The making of a diagnosis of OA where appropriate is regarded as intrinsically important as evidence by its inclusion in the NICE quality standards. It is not known whether the diagnosis itself prompts clinicians to adhere more closely to recommended management strategies, or whether the diagnosis and improved management are both associated with unobserved variables (confounders) in this study. It seems plausible that encouraging clinicians to make and record an accurate diagnosis of OA may prompt better guideline adherence as well as act as a basis for clinical audit. Multiple consultation associated with better indicator achievement was unsurprising, though it is not clear if this reflects only greater opportunity to treat, a more consumerist patient philosophy, or greater disease severity. Given that OA is known to be marginalised in consultations,[62] greater opportunity to manage it proactively might be beneficial, possibly through a structured chronic disease review in combination with other relevant conditions (for example, linked with cardio-metabolic conditions that would also benefit from a focus on weight and exercise). The link with clinician index consultation count would warrant further exploration. It is possible that some clinicians have either a low OA workload, or that OA consultations with those clinicians go frequently uncoded and are less concordant with guidelines. Feedback and education, possibly coupled with structured patient recall systems, may improve care for these clinicians' patients. Given the associations between greater morbidity levels (higher drug chapter counts) and reduced odds of recorded pain and function assessment and exercise advice, it is possible that a greater emphasis on non-pharmacological management may help rebalance assessment and management to be more consistent with the NICE guidelines.

9.3.4 Quality indicators as trial outcomes

A potentially important aspect of quality indicators as trial outcomes, such as in cluster trials, is the ability to use larger samples of patients and the reduction of the potential for selection bias in trial recruitment. Attrition during the trial would also be less likely as the outcomes depend on data collected through routine clinical contacts. The indicators demonstrated some capability to discriminate between the intervention and control trial arms, relating to the provision of written information and physiotherapy referral. Lack of discrimination on the basis of other indicators may have been related to the variation in baseline performance between trial arms.

Where there is a strong link between aspects of care suitable for quality indicators and relevant trial outcomes, it seems that their use as outcome measures in place of other more traditional measures may be appropriate. In a study protocol for a cluster trial of implementation of international OA treatment guidelines,[348] Østerås et al. have specified the use of patient self-reported OA quality indicators (which were not dissimilar to those used in MOSAICS) as the primary outcome. However, the most feasible indicators would be those that can be measured through the routine medical record, since bespoke recording systems appear to have some effect on clinical practice that may make determination of the effects of another intervention more complex (by inflating baseline achievement and apparently weakening any independent effect of another intervention). As discussed in Chapter Three, section 3.8, a smartphone application to facilitate writing a patient's own data to the EHR would be one way to facilitate this.

9.4 Proposals for implementation

From the systematic review, a need for structured implementation of the quality indicators had need identified. The indicators derived from this review have been used in the development of patient-reported OA quality indicators.[349]

I had originally envisaged that the result of the research work would be a good candidate for a future iteration of the QOF,[170] especially supplemented by the work to define clinical OA in general practice records. However, there have been major criticisms of the QOF as a vehicle for delivering

improved care in general practice[350-352] and, with the reducing number of points available for quality work contained within the QOF,[170] this now seems a less feasible strategy for implementation.

For OA, the core aspects of management overlap heavily with those for other conditions (education, exercise, and weight loss). With the increasingly recognised and burgeoning public health problems of inactivity and obesity,[353] there is a need across morbidities to promote interventions that tackle these problems. It is the role of NICE and commissioners of health and public health services to determine which interventions are clinically and cost-effective and for clinicians to participate appropriately in signposting to or delivering these: quality indicators can play a useful role in recording which patients with a capacity to benefit have been referred and facilitating peer-referencing and audit.

Since the publication of the systematic review paper[107] and the paper that describes the effect of the template on recorded quality of care,[99] other routes to implementation have opened. The NICE (2014) OA management guidelines reference the systematic review findings as one example of research impact. There is a substantial overlap between the quality indicators that resulted from the systematic review in Chapter Two and the NICE Quality Standards for the care of OA,[334] relating to assessment of pain and function, exercise, weight loss, and use of core interventions before referral for consideration of surgery. One of the NICE-endorsed resources to support the OA quality standards is the template described in Chapter Three, section 3.4.[††] The extent of update of the NICE OA quality standards in primary care is not known. NICE maintains a list of uptake data[354] and as at 08/09/2016, OA did not appear in the list of uptake data. However, for many of the other clinical conditions, the level of uptake is unclear (due to missing data) or suboptimal.

The findings of the research into the effects of the template as described in Chapter Seven, coupled with OA as a traditionally neglected condition in general practice and its status as a condition causing substantial secondary care expenditure, have been enough for all but one MOSAICS

[††]For more information about the template as a NICE-endorsed resource, see https://www.nice.org.uk/guidance/qs87/resources/endorsed-resources-the-osteoarthritis-etemplate-552602701 [accessed 25/08/2016]

practices to wish to continue with the use of the template. Additionally, a local Clinical Commissioning Group (CCG) has developed a Local Enhanced Service for GPs in fifteen pilot practices to participate in improving care for OA, based on (i) model OA consultation training and support materials (ii) use of the indicators and (iii) template with an associated programme of data collection with audit and reflection. This local roll-out of the model OA consultation with template is referred to as Joint Implementation of Guidelines for oSteoArthritis in the West Midlands (JIGSAW).

Internationally, Professor Dziedzic has led a modification of JIGSAW, referred to as JIGSAW-E, which has been granted funding from the European Union via the European Institute of Innovation and Technology[‡‡] (EIT-Health) to collaborate on the implementation of a similar system of enhanced OA care in Denmark, the Netherlands, Norway, and Portugal. This will include local adaptations of the template to prompt and record OA care. The candidate is a member of the JIGSAW and JIGSAW-E teams, advising on template implementation and evaluation.

9.5 Research implications

Quality indicators have an established use in medical audit and quality improvement programmes as a means of testing the success of other quality improvement initiatives rather than a means of quality improvement in their own right. The extent to which their use causes an improvement in quality of care independently of other initiatives is unclear.

Variation in quality of care is a priority for policy makers who see unwarranted variation as an indication of a need for quality improvement as well as a cause of excess expenditure; reminder systems have been identified as one intervention to reduce variation.[355] The extent of baseline variation seen in this analysis implies that much work remains to be done in identifying and reducing variation in quality of care. The template-derived indicators were associated with a relatively substantial degree of variation explained at the level of the clinician, by unobserved factors. There

[‡‡]For further information about the work of EIT-Health, see https://eit.europa.eu/eit-community/eit-health [accessed 24/08/2016]

is also a need to improve adherence to guidelines more generally, for example in the use of X-rays, since their use is unnecessary for diagnosis and their use in management decisions is also not recommended due to poor linkage between symptoms and radiographic changes.

Further work is needed to understand how well recorded quality indicators reflect the actual level of quality of care delivered, and that experienced by the patient. Also, the reduced frequency of template completion in the trial period suggests that the template may not be durable as a means of measuring and changing clinician behaviour (with respect to prescribing and weight recording, as noted in Chapter Seven). The causes for this reduction should be investigated in further work to determine whether there are adjustments that maintain the usefulness of the template as a clinical prompt and as a means for collecting data for subsequent clinical audit.

Given the known problems with pharmacological management of OA (drug effectiveness and side effects,[139,356] and the lack of disease-modifying agents[357]), it is disappointing that there was not more of a recorded focus on the non-pharmacological management of OA. Strategies known to be potentially effective (exercise, predominantly, supplemented with the other core interventions of education and weight loss where appropriate) have been shown not to be consistently recommended and recorded within the primary care record. Continued efforts to identify effective implementation of non-pharmacological management strategies within primary care are needed.

9.6 Conclusion

Despite the availability of suitable quality indicators for OA in general practice, only a limited range of aspects of quality of care can be measured through routinely-recorded information. This suggests that the longstanding arguments for better data quality and quantity to enable assessment of the quality of care have yet to make an impact in OA care.

Although there is variation in recorded quality of care between clinicians, templates such as that used in this study hold promise for enhancement of the routine EMR to facilitate the primary care of OA though information continuity and to form the basis of medical audit and other quality

improvement initiatives that require continuous data collection. Data collected in this way is capable of detecting change in aspects of OA care over time.

In the future, two particular aspects of quality should be explored further: the variation between clinicians in recorded care and engagement with quality improvement initiatives needs to be understood and reduced, and the link between quality of care as recorded in the medical record and experienced by patients requires further investigation.

References

1. Dziedzic KS, Healey EL, Porcheret M*, et al.* Implementing the NICE osteoarthritis guidelines: a mixed methods study and cluster randomised trial of a model osteoarthritis consultation in primary care--the Management of OsteoArthritis In Consultations (MOSAICS) study protocol. *Implement Sci* 2014;9(1):95 doi: 10.1186/s13012-014-0095-y

2. Dieppe P. The management of osteoarthritis in the third millennium. *Scand J Rheumatol* 2000;29(5):279-81

3. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;16(4):494-502

4. National Collaborating Centre for Chronic Conditions. *Osteoarthritis: national clinical guideline for care and management in adults.* London: Royal College of Physicians, 2008.

5. National Institute for Health & Clinical Excellence. *NICE clinical guideline [CG59] Osteoarthritis: the care and management of osteoarthritis in adults*. London: National Institute for Health & Clinical Excellence, 2008.

6. Schiphof D, de Klerk BM, Koes BW*, et al.* Good reliability, questionable validity of 25 different classification criteria of knee osteoarthritis: a systematic appraisal. *J Clin Epidemiol* 2008;61(12):1205-15 doi: 10.1016/j.jclinepi.2008.04.003

7. Croft P, Cooper C, Wickham C*, et al.* Defining osteoarthritis of the hip for epidemiologic studies. *Am J Epidemiol* 1990;132(3):514-22

8. Radiography Working Group of the OARSI-OMERACT Imaging Workshop, Le Graverand MP, Mazzuca S*, et al.* Assessment of the radioanatomic positioning of the osteoarthritic knee in serial radiographs: comparison of three acquisition techniques. *Osteoarthritis Cartilage* 2006;14 Suppl A:A37-43 doi: 10.1016/j.joca.2006.02.024

9. Duncan RC, Hay EM, Saklatvala J*, et al.* Prevalence of radiographic osteoarthritis - it all depends on your point of view. *Rheumatology (Oxford)* 2006;45(6):757-60 doi: 10.1093/rheumatology/kei270

10. Hunter DJ, Guermazi A, Roemer F*, et al.* Structural correlates of pain in joints with osteoarthritis. *Osteoarthritis Cartilage* 2013;21(9):1170-8 doi: 10.1016/j.joca.2013.05.017

11. Zhang W, Doherty M, Peat G*, et al.* EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis* 2010;69(3):483-9 doi: 10.1136/ard.2009.113100

12. Bedson J, Jordan K, Croft P. How do GPs use *x* rays to manage chronic knee pain in the elderly? A case study. *Ann Rheum Dis* 2003;62(5):450-4

13. Bedson J, Jordan KP, Croft PR. A cross sectional study of requests for knee radiographs from primary care. *BMC Musculoskelet Disord* 2007;8(1):77 doi: 10.1186/1471-2474-8-77

14. Kim C, Nevitt MC, Niu J*, et al.* Association of hip pain with radiographic evidence of hip osteoarthritis: diagnostic test study. *BMJ* 2015;351 doi: 10.1136/bmj.h5983

15. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskelet Disord* 2008;9(1):116 doi: 10.1186/1471-2474-9-116

16. Felson DT, Lawrence RC, Dieppe PA*, et al.* Osteoarthritis: new insights. Part 1: the disease and its risk factors. *Ann Intern Med* 2000;133(8):635-46 doi: 10.7326/0003-4819-133-8-200010170-00016

17. Jordan KP, Tan V, Edwards JJ*, et al.* Influences on the decision to use an osteoarthritis diagnosis in primary care: a cohort study with linked survey and electronic health record data. *Osteoarthritis Cartilage* 2016;24(5):786-93 doi: 10.1016/j.joca.2015.12.015

18. Lane NE, Brandt K, Hawker G*, et al.* OARSI-FDA initiative: defining the disease state of osteoarthritis. *Osteoarthritis Cartilage* 2011;19(5):478-82 doi: 10.1016/j.joca.2010.09.013

19. Pereira D, Peleteiro B, Araujo J*, et al.* The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. *Osteoarthritis Cartilage* 2011;19(11):1270-85 doi: 10.1016/j.joca.2011.08.009

20. Thomas E, Peat G, Croft P. Defining and mapping the person with osteoarthritis for population studies and public health. *Rheumatology (Oxford)* 2014;53(2):338-45 doi: 10.1093/rheumatology/ket346

21. Jordan K, Clarke AM, Symmons DP*, et al.* Measuring disease prevalence: a comparison of musculoskeletal disease using four general practice consultation databases. *Br J Gen Pract* 2007;57(534):7-14

22. Porcheret M, Hughes R, Evans D*, et al.* Data quality of general practice electronic health records: the impact of a program of assessments, feedback, and training. *J Am Med Inform Assoc* 2004;11(1):78-86 doi: 10.1197/jamia.M1362

23. Jordan KP, Joud A, Bergknut C*, et al.* International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. *Ann Rheum Dis* 2014;73(1):212-8 doi: 10.1136/annrheumdis-2012-202634

24. Hoy DG, Smith E, Cross M*, et al.* The global burden of musculoskeletal conditions for 2010: an overview of methods. *Ann Rheum Dis* 2014;73(6):982-9 doi: 10.1136/annrheumdis-2013-204344

25. Busija L, Bridgett L, Williams SR*, et al.* Osteoarthritis. *Best Pract Res Clin Rheumatol* 2010;24(6):757-68 doi: 10.1016/j.berh.2010.11.001

26. Vos T, Flaxman AD, Naghavi M*, et al.* Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380(9859):2163-96 doi: 10.1016/S0140-6736(12)61729-2

27. Murray CJ, Richards MA, Newton JN*, et al.* UK health performance: findings of the Global Burden of Disease Study 2010. *Lancet* 2013;381(9871):997-1020 doi: 10.1016/S0140-6736(13)60355-4

28. Arthritis Care. *OA Nation 2012*. London, Arthritis Care, 2012.

29. Puig-Junoy J, Ruiz Zamora A. Socio-economic costs of osteoarthritis: a systematic review of cost-of-illness studies. *Semin Arthritis Rheum* 2015;44(5):531-41 doi: 10.1016/j.semarthrit.2014.10.012

30. Bijlsma JW, Knahr K. Strategies for the prevention and management of osteoarthritis of the hip and knee. *Best Pract Res Clin Rheumatol* 2007;21(1):59-76 doi: 10.1016/j.berh.2006.08.013

31. Jordan KM, Arden NK, Doherty M*, et al.* EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Ann Rheum Dis* 2003;62(12):1145-55 doi: 10.1136/ard.2003.011742

32. Zhang W, Doherty M, Arden N*, et al.* EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2005;64(5):669-81 doi: 10.1136/ard.2004.028886

33. Zhang W, Doherty M, Leeb BF*, et al.* EULAR evidence based recommendations for the management of hand osteoarthritis: report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007;66(3):377-88 doi: 10.1136/ard.2006.062091

34. Fernandes L, Hagen KB, Bijlsma JW*, et al.* EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. *Ann Rheum Dis* 2013;72(7):1125-35 doi: 10.1136/annrheumdis-2012-202745

35. Zhang W, Moskowitz RW, Nuki G*, et al.* OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage* 2008;16(2):137-62 doi: 10.1016/j.joca.2007.12.013

36. Zhang W, Moskowitz RW, Nuki G*, et al.* OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthritis Cartilage* 2007;15(9):981-1000 doi: 10.1016/j.joca.2007.06.014

37. Zhang W, Nuki G, Moskowitz RW*, et al.* OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage* 2010;18(4):476-99 doi: 10.1016/j.joca.2010.01.013

38. Hochberg MC, Altman RD, April KT*, et al.* American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res (Hoboken)* 2012;64(4):465-74

39. National Institute for Health & Care Excellence. *NICE clinical guideline [CG177] Osteoarthritis: Care and management in adults*. London: National Institute for Health & Care Excellence, 2014.

40. Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;44(3):Suppl:166-206

41. Donabedian A. The quality of care: how can it be assessed? *JAMA* 1988;260(12):1743-48

42. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open* 2013;3(1) doi: 10.1136/bmjopen-2012-001570

43. Maxwell RJ. Quality assessment in health. *BMJ* 1984;288(6428):1470-2

44. Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. *Int J Qual Health Care* 2000;12(4):281-95

45. McAuliffe WE. Measuring the quality of medical care: process versus outcome. *Milbank Mem Fund Q Health Soc* 1979;57(1):118-52 doi: 10.2307/3349751

46. Brook RH, McGlynn EA, Cleary PD. Quality of health care. Part 2: measuring quality of care. *N Engl J Med* 1996;335(13):966-70 doi: 10.1056/NEJM199609263351311

47. Mant J. Process versus outcome indicators in the assessment of quality of health care. *Int J Qual Health Care* 2001;13(6):475-80 doi: 10.1093/intqhc/13.6.475

48. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. *Int J Qual Health Care* 2001;13(6):469-74

49. Aguado A, Guino E, Mukherjee B*, et al.* Variability in prescription drug expenditures explained by adjusted clinical groups (ACG) case-mix: a cross-sectional study of patient electronic records in primary care. *BMC Health Serv Res* 2008;8:53 doi: 10.1186/1472-6963-8-53

50. Salem-Schatz S, Moore G, Rucker M*, et al.* The case for case-mix adjustment in practice profiling: when good apples look bad. *JAMA* 1994;272(11):871-74

51. Sullivan CO, Omar RZ, Ambler G*, et al.* Case-mix and variation in specialist referrals in general practice. *Br J Gen Pract* 2005;55(516):529-33

52. NHS Institute for Innovation and Improvement. *Quality and service improvement tools: Plan, Do, Study, Act* 2008. http://www.institute.nhs.uk/quality_and_service_improvement_tools/quality_and_service_improvement_tools/plan_do_study_act.html.

53. Gillam S, Siriwardena AN. Frameworks for improvement: clinical audit, the plan-do-study-act cycle and significant event audit. *Qual Prim Care* 2013;21(2):123-30

54. Lawrence M, Olesen F. Indicators of quality in health care. *Eur J Gen Pract* 1997;3(3):103-08

55. Campbell SM, Braspenning J, Hutchinson A*, et al.* Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002;11(4):358-64 doi: 10.1136/qhc.11.4.358

56. Maxwell RJ. Dimensions of quality revisited: from thought to action. *Qual Health Care* 1992;1(3):171-7 doi: 10.1136/qshc.1.3.171

57. Gotzsche PC, Liberati A, Torri V*, et al.* Beware of surrogate outcome measures. *Int J Technol Assess Health Care* 1996;12(2):238-46 doi: 10.1017/S0266462300009594

58. Mainz J. Defining and classifying clinical indicators for quality improvement. *Int J Qual Health Care* 2003;15(6):523-30 doi: 10.1093/intqhc/mzg081

59. Mainz J. Developing evidence-based clinical indicators: a state of the art methods primer. *Int J Qual Health Care* 2003;15 Suppl 1:i5-11

60. Stelfox HT, Straus SE. Measuring quality of care: considering conceptual approaches to quality indicator development and evaluation. *J Clin Epidemiol* 2013;66(12):1328-37 doi: 10.1016/j.jclinepi.2013.05.017

61. Wollersheim H, Hermens R, Hulscher M*, et al.* Clinical indicators: development and applications. *Neth J Med* 2007;65(1):15-22

62. Paskins Z, Sanders T, Croft PR*, et al.* The Identity Crisis of Osteoarthritis in General Practice: A Qualitative Study Using Video-Stimulated Recall. *Ann Fam Med* 2015;13(6):537-44 doi: 10.1370/afm.1866

63. Paskins Z, Sanders T, Hassell AB. Comparison of patient experiences of the osteoarthritis consultation with GP attitudes and beliefs to OA: a narrative review. *BMC Fam Pract* 2014;15:46 doi: 10.1186/1471-2296-15-46

64. Ajzen I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 1991;50(2):179-211 doi: 10.1016/0749-5978(91)90020-t

65. Conner M, Armitage CJ. Extending the Theory of Planned Behavior: A Review and Avenues for Further Research. *Journal of Applied Social Psychology* 1998;28(15):1429-64 doi: 10.1111/j.1559-1816.1998.tb01685.x

66. Porcheret M, Jordan K, Jinks C*, et al.* Primary care treatment of knee pain--a survey in older adults. *Rheumatology (Oxford)* 2007;46(11):1694-700 doi: 10.1093/rheumatology/kem232

67. Kirk SA, Campbell SM, Kennell-Webb S*, et al.* Assessing the quality of care of multiple conditions in general practice: practical and methodological problems. *Qual Saf Health Care* 2003;12(6):421-7 doi: 10.1136/qhc.12.6.421

68. Steel N, Maisey S, Clark A*, et al.* Quality of clinical primary care and targeted incentive payments: an observational study. *Br J Gen Pract* 2007;57(539):449-54

69. Broadbent J, Maisey S, Holland R*, et al.* Recorded quality of primary care for osteoarthritis: an observational study. *Br J Gen Pract* 2008;58(557):839-43 doi: 10.3399/bjgp08X376177

70. Steel N, Bachmann M, Maisey S*, et al.* Self reported receipt of care consistent with 32 quality indicators: national population survey of adults aged 50 or more in England. *BMJ* 2008;337:a957 doi: 10.1136/bmj.a957

71. Asch SM, McGlynn EA, Hogan MM*, et al.* Comparison of quality of care for patients in the Veterans Health Administration and patients in a national sample. *Ann Intern Med* 2004;141(12):938-45

72. Cadogan MP, Schnelle JF, Al-Sammarrai NR*, et al.* A standardized quality assessment system to evaluate pain detection and management in the nursing home. *J Am Med Dir Assoc* 2005;6(1):1-9 doi: 10.1016/j.jamda.2004.12.002

73. Chodosh J, Solomon DH, Roth CP*, et al.* The quality of medical care provided to vulnerable older patients with chronic pain. *J Am Geriatr Soc* 2004;52(5):756-61 doi: 10.1111/j.1532-5415.2004.52214.x

74. Ganz DA, Chang JT, Roth CP*, et al.* Quality of osteoarthritis care for community-dwelling older adults. *Arthritis Rheum* 2006;55(2):241-7 doi: 10.1002/art.21844

75. Higashi T, Shekelle PG, Solomon DH*, et al.* The quality of pharmacologic care for vulnerable older patients. *Ann Intern Med* 2004;140(9):714-20

76. MacLean CH, Louie R, Shekelle PG*, et al.* Comparison of administrative data and medical records to measure the quality of medical care provided to vulnerable older patients. *Med Care* 2006;44(2):141-8

77. McGlynn EA, Asch SM, Adams J*, et al.* The quality of health care delivered to adults in the United States. *N Engl J Med* 2003;348(26):2635-45 doi: 10.1056/NEJMsa022615

78. Wenger NS, Solomon DH, Roth CP*, et al.* The quality of medical care provided to vulnerable community-dwelling older patients. *Ann Intern Med* 2003;139(9):740-7

79. Zingmond DS, Saliba D, Wilber KH*, et al.* Measuring the quality of care provided to dually enrolled Medicare and Medicaid beneficiaries living in nursing homes. *Med Care* 2009;47(5):536-44 doi: 10.1097/MLR.0b013e318190cd8b

80. Li LC, Sayre EC, Kopec JA*, et al.* Quality of nonpharmacological care in the community for people with knee and hip osteoarthritis. *J Rheumatol* 2011;38(10):2230-7 doi: 10.3899/jrheum.110264

81. Runciman WB, Hunt TD, Hannaford NA*, et al.* CareTrack: assessing the appropriateness of health care delivery in Australia. *Med J Aust* 2012;197(2):100-5

82. Østerås N, Garratt A, Grotle M*, et al.* Patient-reported quality of care for osteoarthritis: development and testing of the osteoarthritis quality indicator questionnaire. *Arthritis Care Res (Hoboken)* 2013;65(7):1043-51 doi: 10.1002/acr.21976

83. Chevalier X, Marre JP, de Butler J*, et al.* Questionnaire survey of management and prescription of general practitioners in knee osteoarthritis: a comparison with 2000 EULAR recommendations. *Clin Exp Rheumatol* 2004;22(2):205-12

84. DeHaan MN, Guzman J, Bayley MT*, et al.* Knee osteoarthritis clinical practice guidelines -- how are we doing? *Journal of Rheumatology* 2007;34(10):2099

85. Sarzi-Puttini P, Cimmino MA, Scarpa R*, et al.* Do physicians treat symptomatic osteoarthritis patients properly? Results of the AMICA experience. *Semin Arthritis Rheum* 2005;35(1 Suppl 1):38-42 doi: 10.1016/j.semarthrit.2005.02.005

86. Conrozier T, Marre JP, Payen-Champenois C*, et al.* National survey on the non-pharmacological modalities prescribed by French general practitioners in the treatment of lower limb (knee and hip) osteoarthritis. Adherence to the EULAR recommendations and factors influencing adherence. *Clin Exp Rheumatol* 2008;26(5):793-8

87. Denoeud L, Mazieres B, Payen-Champenois C*, et al.* First line treatment of knee osteoarthritis in outpatients in France: adherence to the EULAR 2000 recommendations and factors influencing adherence. *Ann Rheum Dis* 2005;64(1):70-4 doi: 10.1136/ard.2003.015263

88. Maserejian NN, Fischer MA, Trachtenberg FL*, et al.* Variations among primary care physicians in exercise advice, imaging, and analgesics for musculoskeletal pain: results from a factorial experiment. *Arthritis Care Res (Hoboken)* 2014;66(1):147-56 doi: 10.1002/acr.22143

89. NHS Centre for Reviews & Dissemination. Getting Evidence Into Practice. *Effective Health Care Bulletin* 1999;5(1)

90. Lau R, Stevenson F, Ong BN*, et al.* Achieving change in primary care--effectiveness of strategies for improving implementation of complex interventions: systematic review of reviews. *BMJ Open* 2015;5(12):e009993 doi: 10.1136/bmjopen-2015-009993

91. Brand CA, Ackerman IN, Tropea J. Chronic disease management: improving care for people with osteoarthritis. *Best Pract Res Clin Rheumatol* 2014;28(1):119-42 doi: 10.1016/j.berh.2014.01.011

92. Campbell MK, Piaggio G, Elbourne DR*, et al.* Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661 doi: 10.1136/bmj.e5661

93. Mitchell C, Adebajo A, Hay E*, et al.* Shoulder pain: diagnosis and management in primary care. *BMJ* 2005;331(7525):1124-8 doi: 10.1136/bmj.331.7525.1124

94. Meislin RJ, Sperling JW, Stitik TP. Persistent shoulder pain: epidemiology, pathophysiology, and diagnosis. *Am J Orthop (Belle Mead NJ)* 2005;34(12 Suppl):5-9

95. Walker-Bone K, Reading I, Coggon D*, et al.* The anatomical pattern and determinants of pain in the neck and upper limbs: an epidemiologic study. *Pain* 2004;109(1-2):45-51 doi: 10.1016/j.pain.2004.01.008

96. National Institute for Health & Care Excellence. *CG88: Low back pain in adults: early management* 2009. https://www.nice.org.uk/guidance/CG88.

97. Office for National Statistics. *Population Estimates for UK, England and Wales, Scotland and Northern Ireland*, 2015.

98. Health and Social Care Information Centre. *General and Personal Medical Services (England) 2002-2012*, 2013.

99. Edwards JJ, Jordan KP, Peat G*, et al.* Quality of care for OA: the effect of a point-of-care consultation recording template. *Rheumatology (Oxford)* 2015;**54**(5):844-53 doi: 10.1093/rheumatology/keu411

100. Baker R. Managing quality in primary health care: the need for valid information about performance. *Qual Health Care* 2000;9(2):83-83 doi: 10.1136/qhc.9.2.83

101. NHS Employers. *Investing in general practice – revisions to the GMS contract for 2006–07 in England, stage 1. Annex 7: Specification for a directed enhanced service in England: information management and technology*. London: The NHS Confederation (Employers) Company Ltd, 2005.

102. NHS Employers. *Quality & Outcomes Framework.* 2013. http://www.nhsemployers.org/PayAndContracts/GeneralMedicalServicesContract/QOF/Pages/QualityOutcomesFramework.aspx.

103. Employers N. *Introducing QOF 2004* 2010. http://www.nhsemployers.org/your-workforce/primary-care-contacts/general-medical-services/quality-and-outcomes-framework/introducing-qof-2004.

104. NHS England. Data and Business Rules - Osteoporosis: Secondary Prevention of Fragility Fractures (OST). 2014. http://www.hscic.gov.uk/media/13807/Osteoporosis-Secondary-Prevention-of-Fragility-Fractures-v280/pdf/Osteoporosis_Secondary_Prevention_of_Fragility_Fractures_v28.0.pdf (accessed 08/06/2016).

105. Hochberg MC. Quality measures in osteoarthritis. *Clin Exp Rheumatol* 2007;25(6 Suppl 47):102-6

106. Strombeck B, Petersson IF, Vliet Vlieland TP*, et al.* Health care quality indicators on the management of rheumatoid arthritis and osteoarthritis: a literature review. *Rheumatology (Oxford)* 2013;52(2):382-90 doi: 10.1093/rheumatology/kes266

107. Edwards JJ, Khanna M, Jordan KP*, et al.* Quality indicators for the primary care of osteoarthritis: a systematic review. *Ann Rheum Dis* 2015;74(3):490-8 doi: 10.1136/annrheumdis-2013-203913

108. University of York. Centre for Reviews and Dissemination. *Systematic reviews [electronic resource] : CRD's guidance for undertaking reviews in health care*. 3rd ed. York: Centre for Reviews and Dissemination, 2009.

109. National Institute for Health & Care Excellence. *Evidence search* 2010. https://www.evidence.nhs.uk/.

110. National Library of Medicine. *Medical subject headings* 1999. http://www.nlm.nih.gov/mesh/.

111. Agency for Healthcare Research and Quality. http://www.ahrq.gov.

112. ProQuest. *RefWorks* 2009. https://www.refworks.com/.

113. University of Oxford: Centre for Evidence Based Medicine. *Levels of Evidence* 2009. http://www.cebm.net/index.aspx?o=1025.

114. Campbell SM, Cantrill JA. Consensus methods in prescribing research. *J Clin Pharm Ther* 2001;26(1):5-14

115. Boers M, Brooks P, Strand CV*, et al.* The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998;25(2):198-9

116. Valentini G, Matucci Cerinic M. Disease-specific quality indicators, guidelines and outcome measures in scleroderma. *Clin Exp Rheumatol* 2007;25(6 Suppl 47):159-62

117. Geraedts M, Selbmann HK, Ollenschlaeger G. Critical appraisal of clinical performance measures in Germany. *Int J Qual Health Care* 2003;15(1):79-85 doi: 10.1093/intqhc/15.1.79

118. Mokkink LB, Terwee CB, Patrick DL*, et al.* The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19(4):539-49 doi: 10.1007/s11136-010-9606-8

119. Shekelle PG, MacLean CH, Morton SC*, et al.* Assessing care of vulnerable elders: methods for developing quality indicators. *Ann Intern Med* 2001;135(8 Pt 2):647-52

120. Huff ED. Comprehensive reliability assessment and comparison of quality indicators and their components. *J Clin Epidemiol* 1997;50(12):1395-404 doi: 10.1016/S0895-4356(97)00218-7

121. Moher D, Liberati A, Tetzlaff J*, et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. 2009;339 doi: 10.1136/bmj.b2535

122. ACOVE Investigators. Assessing Care of Vulnerable Elders - 3: Quality Indicators. *Journal of the American Geriatrics Society* 2007;55 Suppl 2:S464

123. Smith KL, Soriano TA, Boal J. Brief communication: National quality-of-care standards in home-based primary care. *Ann Intern Med* 2007;146(3):188-92

124. Steel N, Melzer D, Shekelle PG*, et al.* Developing quality indicators for older adults: transfer from the USA to the UK is feasible. *Qual Saf Health Care* 2004;13(4):260-64 doi: 10.1136/qshc.2004.010280

125. EUMUSC.net. *Health care quality indicators for OA* 2012. http://www.eumusc.net/workpackages_wp6.cfm.

126. Wenger NS, Roth CP, Shekelle P*, et al.* Introduction to the assessing care of vulnerable elders-3 quality indicator measurement set. *J Am Geriatr Soc* 2007;55 Suppl 2:S247-52 doi: 10.1111/j.1532-5415.2007.01328.x

127. MacLean CH, Saag KG, Solomon DH*, et al.* Measuring quality in arthritis care: methods for developing the Arthritis Foundation's quality indicator set. *Arthritis Rheum* 2004;51(2):193-202 doi: 10.1002/art.20248

128. Pencharz JN, MacLean CH. Measuring quality in arthritis care: the Arthritis Foundation's Quality Indicator set for osteoarthritis. *Arthritis Rheum* 2004;51(4):538-48 doi: 10.1002/art.20521

129. American Academy of Orthopaedic Surgeons, Physician Consortium for Performance Improvement. *Osteoarthritis: Physician Performance Measurement Set* 2006. http://www.aaos.org/research/committee/evidence/PCPI_Measures_Osteoarthritis.pdf.

130. Campbell S, Hacker J, Roland M, editors. *Quality Indicators for General Practice*. 1 ed. London: RSM Press, 2002.

131. Marshall M, Campbell S, Hacker J*, et al.*, editors. *Quality Indicators for General Practice*. London: RSM Press, 2002.

132. MacLean CH, Pencharz JN, Saag KG. Quality indicators for the care of osteoarthritis in vulnerable elders. *J Am Geriatr Soc* 2007;55 Suppl 2:S383-91 doi: 10.1111/j.1532-5415.2007.01346.x

133. Saliba D, Solomon D, Rubenstein L*, et al.* Quality indicators for the management of medical conditions in nursing home residents. *J Am Med Dir Assoc* 2005;6(3 Suppl):S36-48 doi: 10.1016/j.jamda.2005.03.022

134. MacLean CH. Quality indicators for the management of osteoarthritis in vulnerable elders. *Ann Intern Med* 2001;135(8 Pt 2):711-21

135. Kerr E, Asch S, Hamilton E*, et al. Quality of Care for General Medical Conditions: A Review of the Literature and Quality Indicators*. Santa Monica, CA: RAND Health, 2000.

136. Higashi T, Shekelle PG, Adams JL*, et al.* Quality of care is associated with survival in vulnerable older patients. *Ann Intern Med* 2005;143(4):274-81

137. Etzioni S, Chodosh J, Ferrell BA*, et al.* Quality indicators for pain management in vulnerable elders. *J Am Geriatr Soc* 2007;55 Suppl 2:S403-8 doi: 10.1111/j.1532-5415.2007.01348.x

138. Shrank WH, Polinski JM, Avorn J. Quality indicators for medication use in vulnerable elders. *J Am Geriatr Soc* 2007;55 Suppl 2:S373-82 doi: 10.1111/j.1532-5415.2007.01345.x

139. Roberts E, Delgado Nunes V, Buckner S*, et al.* Paracetamol: not as safe as we thought? A systematic literature review of observational studies. *Ann Rheum Dis* 2016;75(3):552-9 doi: 10.1136/annrheumdis-2014-206914

140. Trelle S, Reichenbach S, Wandel S*, et al.* Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis. *BMJ* 2011;342:c7086 doi: 10.1136/bmj.c7086

141. National Institute for Health & Care Excellence. *NICE quality standards* 2013. http://www.nice.org.uk/guidance/qualitystandards/qualitystandards.jsp.

142. Institute of Medicine Committee on Future Directions for the National Healthcare Quality and Disparities Reports. *Appendix F, The Expected Population Value of Quality Indicator Reporting (EPV-QIR): A Framework for Prioritizing Healthcare Performance Measurement.* In: Meltzer DO, Chung JW, eds. Washington, DC: National Academies Press (US), 2010.

143. Jones P, Shepherd M, Wells S*, et al.* Review article: what makes a good healthcare quality indicator? A systematic review and validation study. *Emerg Med Australas* 2014;26(2):113-24 doi: 10.1111/1742-6723.12195

144. Marshall MN. Can health care quality indicators be transferred between countries? *Qual Saf Health Care* 2003;12(1):8-12 doi: 10.1136/qhc.12.1.8

145. de Lusignan S, Hague N, van Vlymen J*, et al.* Routinely-collected general practice data are complex, but with systematic processing can be used for quality improvement and research. *Inform Prim Care* 2006;14(1):59-66

146. Health and Social Care Information Centre. *Read Codes* 2014. http://systems.hscic.gov.uk/data/uktc/readcodes.

147. UK Department of Health, Royal College of General Practitioners, British Medical Association. The Good Practice Guidelines for GP electronic patient records v4. 2011. http://www.dh.gov.uk/publications.

148. Clinical Terminology Browser [program]. Version 1.04: NHS Information Authority, 2013.

149. de Lusignan S. The barriers to clinical coding in general practice: a literature review. *Med Inform Internet Med* 2005;30(2):89-97 doi: 10.1080/14639230500298651

150. Swinglehurst D, Greenhalgh T, Roberts C. Computer templates in chronic disease management: ethnographic case study in general practice. *BMJ Open* 2012;2(6) doi: 10.1136/bmjopen-2012-001754

151. Black AD, Car J, Pagliari C*, et al.* The impact of eHealth on the quality and safety of health care: a systematic overview. *PLoS Med* 2011;8(1):e1000387 doi: 10.1371/journal.pmed.1000387

152. Shojania KG, Jennings A, Mayhew A*, et al.* The effects of on-screen, point of care computer reminders on processes and outcomes of care. *Cochrane Database Syst Rev* 2009(3):CD001096 doi: 10.1002/14651858.CD001096.pub2

153. Lecumberri R, Marques M, Diaz-Navarlaz MT*, et al.* Maintained effectiveness of an electronic alert system to prevent venous thromboembolism among hospitalized patients. *Thrombosis and Haemostasis* 2008;100(4):699-704 doi: 10.1160/Th08-05-0337

154. Tamblyn R, Huang A, Taylor L*, et al.* A randomized trial of the effectiveness of on-demand versus computer-triggered drug decision support in primary care. *J Am Med Inform Assoc* 2008;15(4):430-8 doi: 10.1197/jamia.M2606

155. Lo HG, Matheny ME, Seger DL*, et al.* Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care. *J Am Med Inform Assoc* 2009;16(1):66-71 doi: 10.1197/jamia.M2687

156. Sundaram V, Lazzeroni LC, Douglass LR*, et al.* A randomized trial of computer-based reminders and audit and feedback to improve HIV screening in a primary care setting. *Int J STD AIDS* 2009;20(8):527-33 doi: 10.1258/ijsa.2008.008423

157. Walker J, Fairley CK, Walker SM*, et al.* Computer reminders for Chlamydia screening in general practice: a randomized controlled trial. *Sex Transm Dis* 2010;37(7):445-50 doi: 10.1097/OLQ.0b013e3181cfcb4c

158. Holt TA, Thorogood M, Griffiths F. Changing clinical practice through patient specific reminders available at the time of the clinical encounter: systematic review and meta-analysis. *J Gen Intern Med* 2012;27(8):974-84 doi: 10.1007/s11606-012-2025-5

159. Robbins GK, Lester W, Johnson KL*, et al.* Efficacy of a clinical decision-support system in an HIV practice: a randomized trial. *Ann Intern Med* 2012;157(11):757-66 doi: 10.7326/0003-4819-157-11-201212040-00003

160. Were MC, Nyandiko WM, Huang KT*, et al.* Computer-generated reminders and quality of pediatric HIV care in a resource-limited setting. *Pediatrics* 2013;131(3):e789-96 doi: 10.1542/peds.2012-2072

161. Kortteisto T, Raitanen J, Komulainen J*, et al.* Patient-specific computer-based decision support in primary healthcare--a randomized trial. *Implement Sci* 2014;9:15 doi: 10.1186/1748-5908-9-15

162. Cheung A, Weir M, Mayhew A*, et al.* Overview of systematic reviews of the effectiveness of reminders in improving healthcare professional behavior. *Syst Rev* 2012;1(1):36 doi: 10.1186/2046-4053-1-36

163. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med* 2009;6(5):e1000065 doi: 10.1371/journal.pmed.1000065

164. Bates DW, Kuperman GJ, Wang S*, et al.* Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10(6):523-30 doi: 10.1197/jamia.M1370

165. Krebs EE, Carey TS, Weinberger M. Accuracy of the pain numeric rating scale as a screening test in primary care. *J Gen Intern Med* 2007;22(10):1453-8 doi: 10.1007/s11606-007-0321-2

166. Lorenz KA, Sherbourne CD, Shugarman LR*, et al.* How reliable is pain as the fifth vital sign? *J Am Board Fam Med* 2009;22(3):291-8 doi: 10.3122/jabfm.2009.03.080162

167. Goulet JL, Brandt C, Crystal S*, et al.* Agreement between electronic medical record-based and self-administered pain numeric rating scale: clinical and research implications. *Med Care* 2013;51(3):245-50 doi: 10.1097/MLR.0b013e318277f1ad

168. Arthritis Research UK. *Exercise and arthritis*. http://www.arthritisresearchuk.org/arthritis-information/arthritis-and-daily-life/exercise-and-arthritis.aspx.

169. Foster NE, Williams B, Grove S*, et al.* The evidence for and against 'PhysioDirect' telephone assessment and advice services. *Physiotherapy* 2011;97(1):78-82 doi: 10.1016/j.physio.2010.05.003

170. NHS Employers. *Changes to QOF 2014/15* 2013. http://www.nhsemployers.org/PAYANDCONTRACTS/GENERALMEDICALSERVICESCONTRACT/QOF/Pages/ChangestoQOF201415.aspx.

171. Hersh WR, Weiner MG, Embi PJ*, et al.* Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med Care* 2013;51(8 Suppl 3):S30-7 doi: 10.1097/MLR.0b013e31829b1dbd

172. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. *Ann Rheum Dis* 2001;60(2):91-7

173. Duncan R, Peat G, Thomas E*, et al.* Symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? *Ann Rheum Dis* 2007;66(1):86-91 doi: 10.1136/ard.2006.052548

174. Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health* 2004;58(8):635-41 doi: 10.1136/jech.2003.008466

175. Aaron SD, Vandemheen KL, Boulet LP*, et al.* Overdiagnosis of asthma in obese and nonobese adults. *CMAJ* 2008;179(11):1121-31 doi: 10.1503/cmaj.081332

176. Ward DG, Halpin DM, Seamark DA. How accurate is diagnosis of asthma in a general practice database? A review of patients' notes and questionnaire-reported symptoms. *Br J Gen Pract* 2004;54(507):753-8

177. Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ* 2012;344:e3502 doi: 10.1136/bmj.e3502

178. National Institute for Health & Clinical Excellence. *Clinical Guideline 127: Hypertension - Clinical management of primary hypertension in adults (Update)* 2011. http://guidance.nice.org.uk/CG127.

179. Kendrick T, Stuart B, Newell C*, et al.* Changes in rates of recorded depression in English primary care 2003-2013: Time trend analyses of effects of the economic recession, and the GP contract quality outcomes framework (QOF). *J Affect Disord* 2015;180:68-78 doi: 10.1016/j.jad.2015.03.040

180. Fleming DM. Morbidity registration and the fourth general practice morbidity survey in England and Wales. *Scand J Prim Health Care Suppl* 1993;2(sup2):37-41 doi: 10.3109/02813439309045500

181. Morden A, Ong BN, Brooks L*, et al.* Introducing Evidence Through Research "Push": Using Theory and Qualitative Methods. *Qual Health Res* 2015;25(11):1560-75 doi: 10.1177/1049732315570120

182. Ong BN, Morden A, Brooks L*, et al.* Changing policy and practice: making sense of national guidelines for osteoarthritis. *Soc Sci Med* 2014;106:101-9 doi: 10.1016/j.socscimed.2014.01.036

183. O'Brien MA, Rogers S, Jamtvedt G*, et al.* Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2007(4):CD000409 doi: 10.1002/14651858.CD000409.pub2

184. Ivers NM, Schwalm JD, Grimshaw JM*, et al.* Delayed educational reminders for long-term medication adherence in ST-elevation myocardial infarction (DERLA-STEMI): protocol for a pragmatic, cluster-randomized controlled trial. *Implement Sci* 2012;7:54 doi: 10.1186/1748-5908-7-54

185. Patient.info. *Patient Access - Apple Health app integration* 2016. http://patient.info/personal-health-record.

186. de Lusignan S, Valentin T, Chan T*, et al.* Problems with primary care data quality: osteoporosis as an exemplar. *Inform Prim Care* 2004;12(3):147-56

187. Sethi D, Wheeler J, Rodrigues LC*, et al.* Investigation of under-ascertainment in epidemiological studies based in general practice. *Int J Epidemiol* 1999;28(1):106-12

188. Rollason W, Khunti K, de Lusignan S. Variation in the recording of diabetes diagnostic data in primary care computer systems: implications for the quality of care. *Inform Prim Care* 2009;17(2):113-9

189. NHS Employers & British Medical Association. *Revisions to the GMS contract 2006/07: Delivering investment in general practice*. London: NHS Employers and the General Practitioners Committee, 2006.

190. NHS England. Enhanced Service Specification: Dementia Identification Scheme. 2014. http://www.england.nhs.uk/wp-content/uploads/2014/10/dementia-ident-schm-fin.pdf (accessed 19/02/2015).

191. Fleming DM. The measurement of morbidity in general practice. *J Epidemiol Community Health* 1991;45(3):180-3

192. Jordan K, Porcheret M, Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. *Fam Pract* 2004;21(4):396-412 doi: 10.1093/fampra/cmh409

193. Hollowell J. The General Practice Research Database: quality of morbidity data. *Popul Trends* 1997(87):36-40

194. Bedson J, Jordan K, Croft P. The prevalence and history of knee osteoarthritis in general practice: a case-control study. *Fam Pract* 2005;22(1):103-8 doi: 10.1093/fampra/cmh700

195. Yu D, Peat G, Bedson J*, et al.* Weighted cumulative exposure models helped identify an association between early knee-pain consultations and future knee OA diagnosis. *J Clin Epidemiol* 2016;76:218-28 doi: 10.1016/j.jclinepi.2016.02.025

196. de Jong J, Visser MR, Wieringa-de Waard M. Which barriers affect morbidity registration performance of GP trainees and trainers? *Int J Med Inform* 2013;82(8):708-16 doi: 10.1016/j.ijmedinf.2013.02.002

197. Bertakis KD, Azari R, Callahan EJ. Patient pain in primary care: Factors that influence physician diagnosis. *Ann Fam Med* 2004;2(3):224-30 doi: 10.1370/afm.66

198. Bertakis KD, Callahan EJ, Helms LJ*, et al.* The effect of patient health status on physician practice style. *Fam Med* 1993;25(8):530-5

199. Coxon D, Frisher M, Jinks C*, et al.* The relative importance of perceived doctor's attitude on the decision to consult for symptomatic osteoarthritis: a choice-based conjoint analysis study. *BMJ Open* 2015;5(10):e009625 doi: 10.1136/bmjopen-2015-009625

200. Birtwhistle R, Keshavjee K, Lambert-Lanning A*, et al.* Building a pan-Canadian primary care sentinel surveillance network: initial development and moving forward. *J Am Board Fam Med* 2009;22(4):412-22 doi: 10.3122/jabfm.2009.04.090081

201. Kadhim-Saleh A, Green M, Williamson T*, et al.* Validation of the diagnostic algorithms for 5 chronic conditions in the Canadian Primary Care Sentinel Surveillance Network (CPCSSN): a Kingston Practice-based Research Network (PBRN) report. *J Am Board Fam Med* 2013;26(2):159-67 doi: 10.3122/jabfm.2013.02.120183

202. Williamson T, Green ME, Birtwhistle R*, et al.* Validating the 8 CPCSSN case definitions for chronic disease surveillance in a primary care database of electronic health records. *Ann Fam Med* 2014;12(4):367-72 doi: 10.1370/afm.1644

203. Coleman N, Halas G, Peeler W*, et al.* From patient care to research: a validation study examining the factors contributing to data quality in a primary care electronic medical record database. *BMC Fam Pract* 2015;16:11 doi: 10.1186/s12875-015-0223-z

204. Harrold LR, Yood RA, Andrade SE*, et al.* Evaluating the predictive value of osteoarthritis diagnoses in an administrative database. *Arthritis Rheum* 2000;43(8):1881-5 doi: 10.1155/2016/6475318

205. Harrold LR, Yood RA, Straus W*, et al.* Challenges of estimating health service utilization for osteoarthritis patients on a population level. *J Rheumatol* 2002;29(9):1931-6

206. Haugen IK, Englund M, Aliabadi P*, et al.* Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. *Ann Rheum Dis* 2011;70(9):1581-6 doi: 10.1136/ard.2011.150078

207. Jordan JM, Helmick CG, Renner JB*, et al.* Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. *J Rheumatol* 2007;34(1):172-80

208. Jordan JM, Helmick CG, Renner JB*, et al.* Prevalence of hip symptoms and radiographic and symptomatic hip osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. *J Rheumatol* 2009;36(4):809-15 doi: 10.3899/jrheum.080677

209. Suri P, Morgenroth DC, Hunter DJ. Epidemiology of osteoarthritis and associated comorbidities. *PM R* 2012;4(5 Suppl):S10-9 doi: 10.1016/j.pmrj.2012.01.007

210. Kopec JA, Rahman MM, Berthelot JM*, et al.* Descriptive epidemiology of osteoarthritis in British Columbia, Canada. *J Rheumatol* 2007;34(2):386-93

211. Visser AW, de Mutsert R, le Cessie S*, et al.* The relative contribution of mechanical stress and systemic processes in different types of osteoarthritis: the NEO study. *Ann Rheum Dis* 2015;74(10):1842-7 doi: 10.1136/annrheumdis-2013-205012

212. Blagojevic M, Jinks C, Jeffery A*, et al.* Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage* 2010;18(1):24-33 doi: 10.1016/j.joca.2009.08.010

213. Sellam J, Berenbaum F. Is osteoarthritis a metabolic disease? *Joint Bone Spine* 2013;80(6):568-73 doi: 10.1016/j.jbspin.2013.09.007

214. Magnusson K, Osteras N, Haugen IK*, et al.* No strong relationship between body mass index and clinical hand osteoarthritis: results from a population-based case-control study. *Scand J Rheumatol* 2014;43(5):409-15 doi: 10.3109/03009742.2014.900700

215. Hoogeboom TJ, den Broeder AA, Swierstra BA*, et al.* Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study. *Arthritis Care Res (Hoboken)* 2012;64(1):54-8 doi: 10.1002/acr.20647

216. Heo M, Faith MS, Mott JW*, et al.* Hierarchical linear models for the development of growth curves: an example with body mass index in overweight/obese adults. *Stat Med* 2003;22(11):1911-42 doi: 10.1002/sim.1218

217. Joint Formulary Committee. *British National Formulary*: Pharmaceutical Press, 2011.

218. Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. *Fam Pract* 2013;30(2):172-8 doi: 10.1093/fampra/cms060

219. Charlson ME, Pompei P, Ales KL*, et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373-83 doi: 10.1016/0021-9681(87)90171-8

220. Kaufman DW, Kelly JP, Rosenberg L*, et al.* Recent patterns of medication use in the ambulatory adult population of the United States - The Slone survey. *JAMA* 2002;287(3):337-44 doi: 10.1001/jama.287.3.337

221. Patterson SM, Hughes C, Kerse N*, et al.* Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database Syst Rev* 2012;5(5):CD008165 doi: 10.1002/14651858.CD008165.pub2

222. Hovstadius B, Hovstadius K, Astrand B*, et al.* Increasing polypharmacy - an individual-based study of the Swedish population 2005-2008. *BMC Clin Pharmacol* 2010;10:16 doi: 10.1186/1472-6904-10-16

223. Olsson IN, Runnamo R, Engfeldt P. Drug treatment in the elderly: an intervention in primary care to enhance prescription quality and quality of life. *Scand J Prim Health Care* 2012;30(1):3-9 doi: 10.3109/02813432.2011.629149

224. Office for National Statistics. *Mortality Statistics: Metadata*. Newport: Office for National Statistics, 2014.

225. Office for National Statistics. *Age standardised mortality rate calculation template using the 2013 ESP* 2013. http://www.ons.gov.uk/ons/guide-method/user-guidance/health-and-life-events/age-standardised-mortality-rate-calculation-template-using-the-2013-esp.xls.

226. European Union. *Revision of the European Standard Population: Report of Eurostat's task force*. Luxembourg: Publications Office of the European Union, 2013.

227. Office for National Statistics. *Annual Mid-year Population Estimates, 2013*. Newport: Office for National Statistics, 2014.

228. Spence D. What happened to the doctor-patient relationship? *BMJ* 2012;344:e4349 doi: 10.1136/bmj.e4349

229. University of Bristol Centre for Multilevel Modelling. *LEMMA (Learning environment for multilevel methodology and applications)* 2002-2014. http://www.cmm.bris.ac.uk/lemma/.

230. SPSS for Windows [program]. Version 21.0. Armonk, NY: IBM Corp, 2012.

231. Microsoft® Excel 2013 [program]. Version 2013. Redmond, Washington: Microsoft Corporation, 2013.

232. Confidence Interval Analysis [program]. Version 2.2.0. Southampton: University of Southampton, 2011.

233. Stata [program]. Version 13.1. College Station, TX: StataCorp LP, 2013.

234. MLwiN [program]. Version 2.34. Bristol: Centre for Multilevel Modelling, University of Bristol, 2015.

235. Leckie G, Charlton C. runmlwin: A Program to Run the MLwiN Multilevel Modeling Software from within Stata. *Journal of Statistical Software* 2013;52(11):1-40

236. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37(5):360-3

237. Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. *JAMA* 2011;305(23):2462-63 doi: 10.1001/jama.2011.822

238. Hochberg MC. Mortality in osteoarthritis. *Clin Exp Rheumatol* 2008;26(5 Suppl 51):S120-4

239. Nuesch E, Dieppe P, Reichenbach S*, et al.* All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ* 2011;342:d1165 doi: 10.1136/bmj.d1165

240. Cooper C, Arden NK. Excess mortality in osteoarthritis. *BMJ* 2011;342:d1407 doi: 10.1136/bmj.d1407

241. Kadam UT, Jordan K, Croft PR. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales. *Ann Rheum Dis* 2004;63(4):408-14 doi: 10.1136/ard.2003.007526

242. Bjerrum L, Sogaard J, Hallas J*, et al.* Polypharmacy: correlations with sex, age and drug regimen. A prescription database study. *Eur J Clin Pharmacol* 1998;54(3):197-202 doi: 10.1007/s002280050445

243. van Saase JL, van Romunde LK, Cats A*, et al.* Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. *Ann Rheum Dis* 1989;48(4):271-80

244. Oliveria SA, Felson DT, Reed JI*, et al.* Incidence of Symptomatic Hand, Hip, and Knee Osteoarthritis among Patients in a Health Maintenance Organization. *Arthritis Rheum* 1995;38(8):1134-41 doi: 10.1002/art.1780380817

245. Lawrence RC, Felson DT, Helmick CG*, et al.* Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum* 2008;58(1):26-35 doi: 10.1002/art.23176

246. Davis MA, Ettinger WH, Neuhaus JM*, et al.* Sex differences in osteoarthritis of the knee. The role of obesity. *Am J Epidemiol* 1988;127(5):1019-30

247. Felson DT, Anderson JJ, Naimark A*, et al.* Obesity and knee osteoarthritis. The Framingham Study. *Ann Intern Med* 1988;109(1):18-24 doi: 10.7326/0003-4819-109-1-18

248. Grotle M, Hagen KB, Natvig B*, et al.* Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. *BMC Musculoskelet Disord* 2008;9:132 doi: 10.1186/1471-2474-9-132

249. Cooper C, Inskip H, Croft P*, et al.* Individual risk factors for hip osteoarthritis: obesity, hip injury, and physical activity. *Am J Epidemiol* 1998;147(6):516-22

250. Tai TW, Anandarajah S, Dhoul N*, et al.* Variation in clinical coding lists in UK general practice: a barrier to consistent data entry? *Inform Prim Care* 2007;15(3):143-50

251. Rothman KJ. Six persistent research misconceptions. *J Gen Intern Med* 2014;29(7):1060-4 doi: 10.1007/s11606-013-2755-z

252. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;1(1):43-46

253. Waring ME, Roberts MB, Parker DR*, et al.* Documentation and management of overweight and obesity in primary care. *J Am Board Fam Med* 2009;22(5):544-52 doi: 10.3122/jabfm.2009.05.080173

254. Hensor EM, Dube B, Kingsbury SR*, et al.* Toward a clinical definition of early osteoarthritis: onset of patient-reported knee pain begins on stairs. Data from the osteoarthritis initiative. *Arthritis Care Res (Hoboken)* 2015;67(1):40-7 doi: 10.1002/acr.22418

255. Khan NF, Perera R, Harper S*, et al.* Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. *BMC Fam Pract* 2010;11:1 doi: 10.1186/1471-2296-11-1

256. Urwin M, Symmons D, Allison T*, et al.* Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. *Ann Rheum Dis* 1998;57(11):649-55 doi: 10.1136/ard.57.11.649

257. Jinks C, Jordan K, Croft P. Disabling knee pain--another consequence of obesity: results from a prospective cohort study. *BMC Public Health* 2006;6:258 doi: 10.1186/1471-2458-6-258

258. Liu B, Balkwill A, Banks E*, et al.* Relationship of height, weight and body mass index to the risk of hip and knee replacements in middle-aged women. *Rheumatology (Oxford)* 2007;46(5):861-7 doi: 10.1093/rheumatology/kel434

259. Harrold LR, Li W, Yood RA*, et al.* Identification of patients with arthritis and arthritis-related functional limitation using administrative data. *Journal of Public Health Management and Practice* 2008;14(5):487-97

260. Fransen M, Simic M, Harmer AR. Determinants of MSK health and disability: lifestyle determinants of symptomatic osteoarthritis. *Best Pract Res Clin Rheumatol* 2014;28(3):435-60 doi: 10.1016/j.berh.2014.07.002

261. Jordan KP, Hayward R, Roberts E*, et al.* The relationship of individual and neighbourhood deprivation with morbidity in older adults: an observational study. *Eur J Public Health* 2014;24(3):396-8 doi: 10.1093/eurpub/ckt160

262. Rolfson O, Wissig S, van Maasakkers L*, et al.* Defining an International Standard Set of Outcome Measures for Patients With Hip or Knee Osteoarthritis: Consensus of the International Consortium for Health Outcomes Measurement Hip and Knee Osteoarthritis Working Group. *Arthritis Care Res (Hoboken)* 2016;68(11):1631-39 doi: 10.1002/acr.22868

263. NHS Employers. *Changes to QOF 2013/14* 2012. http://www.nhsemployers.org/your-workforce/primary-care-contacts/general-medical-services/quality-and-outcomes-framework/changes-to-qof-201314.

264. Ndlovu M, Bedson J, Jones PW*, et al.* Pain medication management of musculoskeletal conditions at first presentation in primary care: analysis of routinely collected medical record data. *BMC Musculoskelet Disord* 2014;15:418 doi: 10.1186/1471-2474-15-418

265. Health and Social Care Information Centre. *Trends in Consultation Rates in General Practice 1995 to 2008: Analysis of the QResearch® Database*: The NHS Information Centre for health and social care, 2009.

266. Cepeda MS, Camargo F, Zea C*, et al.* Tramadol for osteoarthritis. *Cochrane Database Syst Rev* 2006;3(3):CD005522 doi: 10.1002/14651858.CD005522.pub2

267. Nuesch E, Rutjes AW, Husni E*, et al.* Oral or transdermal opioids for osteoarthritis of the knee or hip. *Cochrane Database Syst Rev* 2009;4(4):CD003115 doi: 10.1002/14651858.CD003115.pub3

268. Padwal RS, Majumdar SR. Drug treatments for obesity: orlistat, sibutramine, and rimonabant. *Lancet* 2007;369(9555):71-7 doi: 10.1016/S0140-6736(07)60033-6

269. Wetle T. Age as a risk factor for inadequate treatment. *JAMA* 1987;258(4):516

270. Miller M, Byington R, Hunninghake D*, et al.* Sex bias and underutilization of lipid-lowering therapy in patients with coronary artery disease at academic medical centers in the United States and Canada. Prospective Randomized Evaluation of the Vascular Effects of Norvasc Trial (PREVENT) Investigators. *Arch Intern Med* 2000;160(3):343-7

271. Piette JD, Bibbins-Domingo K, Schillinger D. Health care discrimination, processes of care, and diabetes patients' health status. *Patient Educ Couns* 2006;60(1):41-8 doi: 10.1016/j.pec.2004.12.001

272. Foster GD, Wadden TA, Makris AP*, et al.* Primary care physicians' attitudes about obesity and its treatment. *Obes Res* 2003;11(10):1168-77 doi: 10.1038/oby.2003.161

273. MLwiN [program]. Version 2.1. Bristol: Centre for Multilevel Modelling, University of Bristol, 2009.

274. Townsend A, Wyke S, Hunt K. Frequent consulting and multiple morbidity: a qualitative comparison of 'high' and 'low' consulters of GPs. *Fam Pract* 2008;25(3):168-75 doi: 10.1093/fampra/cmn017

275. Lupton D. Consumerism, reflexivity and the medical encounter. *Soc Sci Med* 1997;45(3):373-81 doi: 10.1016/S0277-9536(96)00353-X

276. Min LC, Wenger NS, Fung C*, et al.* Multimorbidity is associated with better quality of care among vulnerable elders. *Med Care* 2007;45(6):480-8 doi: 10.1097/MLR.0b013e318030fff9

277. Gronhaug G, Osteras N, Hagen KB. Quality of hip and knee osteoarthritis management in primary health care in a Norwegian county: a cross-sectional survey. *BMC Health Serv Res* 2014;14(1):598 doi: 10.1186/s12913-014-0598-x

278. Fraenkel L, Bogardus ST, Jr., Concato J*, et al.* Treatment options in knee osteoarthritis: the patient's perspective. *Arch Intern Med* 2004;164(12):1299-304 doi: 10.1001/archinte.164.12.1299

279. Fransen M, McConnell S, Harmer AR*, et al.* Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2015;1:CD004376 doi: 10.1002/14651858.CD004376.pub3

280. Fransen M, McConnell S, Hernandez-Molina G*, et al.* Exercise for osteoarthritis of the hip. *Cochrane Database Syst Rev* 2014;4(4):CD007912 doi: 10.1002/14651858.CD007912.pub2

281. Memel DS, Kirwan JR, Sharp DJ*, et al.* General practitioners miss disability and anxiety as well as depression in their patients with osteoarthritis. *Br J Gen Pract* 2000;50(457):645-48

282. Stevens A. *Health care needs assessment: the epidemiologically based needs assessment reviews*: Radcliffe Publishing, 2004.

283. Frankel S, Eachus J, Pearson N*, et al.* Population requirement for primary hip-replacement surgery: a cross-sectional study. *Lancet* 1999;353(9161):1304-9 doi: 10.1016/S0140-6736(98)06451-4

284. Judge A, Welton NJ, Sandhu J*, et al.* Equity in access to total joint replacement of the hip and knee in England: cross sectional study. *BMJ* 2010;341:c4092 doi: 10.1136/bmj.c4092

285. Wertheimer AI. The defined daily dose system (DDD) for drug utilization review. *Hosp Pharm* 1986;21(3):233-34, 39-41, 58

286. NHS Connecting for Health. *Choose and Book* 2012. http://www.chooseandbook.nhs.uk/.

287. National Institute for Health Research. *Clinical Research Network* 2014. http://www.crn.nihr.ac.uk.

288. Cottrell E, Roddy E, Foster NE. The attitudes, beliefs and behaviours of GPs regarding exercise for chronic knee pain: a systematic review. *BMC Fam Pract* 2010;11:4 doi: 10.1186/1471-2296-11-4

289. Chartered Society of Physiotherapy. *Stretched to the limit: An audit of physiotherapy services in England*. London: Chartered Society of Physiotherapy, 2012.

290. Young EL, Holt PJ, Poloniecki JD*, et al.* Meta-analysis and systematic review of the relationship between surgeon annual caseload and mortality for elective open abdominal aortic aneurysm repairs. *J Vasc Surg* 2007;46(6):1287-94 doi: 10.1016/j.jvs.2007.06.038

291. Dimick JB, Cowan JA, Jr., Stanley JC*, et al.* Surgeon specialty and provider volumes are related to outcome of intact abdominal aortic aneurysm repair in the United States. *J Vasc Surg* 2003;38(4):739-44 doi: 10.1016/S0741-5214(03)00470-1

292. Pearce WH, Parker MA, Feinglass J*, et al.* The importance of surgeon volume and training in outcomes for vascular surgical procedures. *J Vasc Surg* 1999;29(5):768-76; discussion 77-8 doi: 10.1016/S0741-5214(99)70202-8

293. Ashworth M, Armstrong D. The relationship between general practice characteristics and quality of care: a national survey of quality indicators used in the UK Quality and Outcomes Framework, 2004–5. *BMC Fam Pract* 2006;7(1):68

294. Hallam L, Metcalfe DH. Seasonal variations in the process of care in urban general practice. *J Epidemiol Community Health* 1985;39(1):90-3 doi: 10.1136/jech.39.1.90

295. Waller J, Hodgkin P. *General Practice: Demanding Work: Understanding Patterns of Work in Primary Care*: Radcliffe Publishing, 2000.

296. Jensen OM. *Cancer registration: principles and methods*: IARC, 1991.

297. StatsDirect statistical software [program]. Version 3.0.177. England: StatsDirect Ltd., 2016.

298. Stata [program]. Version 14.1. College Station, TX: StataCorp LP, 2015.

299. Rose EA, Deshikachar AM, Schwartz KL*, et al.* Use of a template to improve documentation and coding. *Fam Med* 2001;33(7):516

300. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002;359(9302):248-52 doi: 10.1016/S0140-6736(02)07451-2

301. Beck CA, Richard H, Tu JV*, et al.* Administrative Data Feedback for Effective Cardiac Treatment: AFFECT, a cluster randomized trial. *JAMA* 2005;294(3):309-17 doi: 10.1001/jama.294.3.309

302. Bardach NS, Wang JJ, De Leon SF*, et al.* Effect of pay-for-performance incentives on quality of care in small practices with electronic health records: a randomized trial. *JAMA* 2013;310(10):1051-9 doi: 10.1001/jama.2013.277353

303. Sherr K, Gimbel S, Rustagi A*, et al.* Systems analysis and improvement to optimize pMTCT (SAIA): a cluster randomized trial. *Implement Sci* 2014;9(1):55 doi: 10.1186/1748-5908-9-55

304. Grime J, Dudley B. Developing written information on osteoarthritis for patients: facilitating user involvement by exposure to qualitative research. *Health Expect* 2014;17(2):164-73 doi: 10.1111/j.1369-7625.2011.00741.x

305. Porcheret M, Grime J, Main C*, et al.* Developing a model osteoarthritis consultation: a Delphi consensus exercise. *BMC Musculoskelet Disord* 2013;14:25 doi: 10.1186/1471-2474-14-25

306. Finney A, Porcheret M, Grime J*, et al.* Defining the content of an opportunistic osteoarthritis consultation with primary health care professionals: a Delphi consensus study. *Arthritis Care Res (Hoboken)* 2013;65(6):962-8 doi: 10.1002/acr.21917

307. Grol R, Wensing M, Eccles M. *Improving patient care: the implementation of change in clinical practice*. Edinburgh, 2005.

308. Improved Clinical Effectiveness through Behavioural Research G. Designing theoretically-informed implementation interventions. *Implement Sci* 2006;1(1):4 doi: 10.1186/1748-5908-1-4

309. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011;6:42 doi: 10.1186/1748-5908-6-42

310. Kennedy A, Rogers A, Chew-Graham C*, et al.* Implementation of a self-management support approach (WISE) across a health system: a process evaluation explaining what did and did not work for organisations, clinicians and patients. *Implement Sci* 2014;9:129 doi: 10.1186/s13012-014-0129-5

311. Jordan K, Jöud A, Bergknut C*, et al.* International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. *Annals of the Rheumatic Diseases* 2013;73(1):212-8 doi: 10.1136/annrheumdis-2012-202634

312. Russell I, Irvine D, Hutchinson A. North of England Study of Standards and Performance in General Practice. Medical audit in general practice: effects on doctors' clinical behaviour and the health of the patients with common childhood conditions. *BMJ* 1992;304:1480-84

313. Adams G, Gulliford MC, Ukoumunne OC*, et al.* Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004;57(8):785-94 doi: 10.1016/j.jclinepi.2003.12.013

314. Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? *BMJ* 2009;339:b4006 doi: 10.1136/bmj.b4006

315. Ivers NM, Halperin IJ, Barnsley J*, et al.* Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. *Trials* 2012;13(1):120 doi: 10.1186/1745-6215-13-120

316. Donner A, Klar N. Methods for comparing event rates in intervention studies when the unit of allocation is a cluster. *Am J Epidemiol* 1994;140(3):279-89; discussion 300-1

317. Østerås N, Jordan KP, Clausen B*, et al.* Self-reported quality care for knee osteoarthritis: comparisons across Denmark, Norway, Portugal and the UK. *RMD Open* 2015;1(1):e000136 doi: 10.1136/rmdopen-2015-000136

318. Eyssen IC, Steultjens MP, de Groot V*, et al.* A cluster randomised controlled trial on the efficacy of client-centred occupational therapy in multiple sclerosis: good process, poor outcome. *Disabil Rehabil* 2013;35(19):1636-46 doi: 10.3109/09638288.2012.748845

319. Frei A, Chmiel C, Schlapfer H*, et al.* The Chronic CARe for diAbeTes study (CARAT): a cluster randomized controlled trial. *Cardiovasc Diabetol* 2010;9:23 doi: 10.1186/1475-2840-9-23

320. Yazdany J, Trupin L, Schmajuk G*, et al.* Quality of care in systemic lupus erythematosus: the association between process and outcome measures in the Lupus Outcomes Study. *BMJ Qual Saf* 2014;23(8):659-66 doi: 10.1136/bmjqs-2013-002494

321. Radestad M, Nilsson H, Castren M*, et al.* Combining performance and outcome indicators can be used in a standardized way: a pilot study of two multidisciplinary, full-scale major aircraft exercises. *Scand J Trauma Resusc Emerg Med* 2012;20:58 doi: 10.1186/1757-7241-20-58

322. Steurer-Stey C, Markun S, Lana KD*, et al.* The improving care in chronic obstructive lung disease study: CAROL improving processes of care and quality of life of COPD patients in primary care: study protocol for a randomized controlled trial. *Trials* 2014;15:96 doi: 10.1186/1745-6215-15-96

323. Maisey S, Steel N, Marsh R*, et al.* Effects of payment for performance in primary care: qualitative interview study. *J Health Serv Res Policy* 2008;13(3):133-9 doi: 10.1258/jhsrp.2008.007118

324. Ashworth M, Kordowicz M. Quality and Outcomes Framework: smoke and mirrors? *Qual Prim Care* 2010;18(2):127-31

325. Ashworth M, Kordowicz M. Quality and Outcomes Framework: time to take stock. *Br J Gen Pract* 2010;60(578):637-8 doi: 10.3399/bjgp10X515313

326. Berkman ND, Sheridan SL, Donahue KE*, et al.* Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med* 2011;155(2):97-107 doi: 10.7326/0003-4819-155-2-201107190-00005

327. Rowlands G, Protheroe J, Winkley J*, et al.* A mismatch between population health literacy and the complexity of health information: an observational study. *Br J Gen Pract* 2015;65(635):e379-86 doi: 10.3399/bjgp15X685285

328. Clarson LE, Nicholl BI, Bishop A*, et al.* Should there be a Quality and Outcomes Framework domain for osteoarthritis? A cross-sectional survey in general practice. *Qual Prim Care* 2013;21(2):97-103

329. Clarson LE, Nicholl BI, Bishop A*, et al.* Monitoring Osteoarthritis: A Cross-sectional Survey in General Practice. *Clin Med Insights Arthritis Musculoskelet Disord* 2013;6:85-91 doi: 10.4137/CMAMD.S12606

330. Salisbury C, Procter S, Stewart K*, et al.* The content of general practice consultations: cross-sectional study based on video recordings. *Br J Gen Pract* 2013;63(616):e751-9 doi: 10.3399/bjgp13X674431

331. Stange KC, Zyzanski SJ, Smith TF*, et al.* How Valid Are Medical Records and Patient Questionnaires for Physician Profiling and Health Services Research?: A Comparison With Direct Observation of Patient Visits. *Med Care* 1998;36(6):851-67

332. Luck J, Peabody JW, Dresselhaus TR*, et al.* How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. *Am J Med* 2000;108(8):642-9

333. Peabody JW, Luck J, Glassman P*, et al.* Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000;283(13):1715-22 doi: 10.1001/jama.283.13.1715

334. National Institute for Health & Care Excellence. *Quality standard for osteoarthritis (NICE quality standard 87)* 2015. https://www.nice.org.uk/guidance/qs87.

335. Sciacovelli L, Sonntag O, Padoan A*, et al.* Monitoring quality indicators in laboratory medicine does not automatically result in quality improvement. *Clin Chem Lab Med* 2012;50(3):463-9 doi: 10.1515/CCLM.2011.809

336. Levene LS, Walker N, Baker R*, et al.* Rationalising data use for general practice: a missed opportunity? *Br J Gen Pract* 2016;66(649):e603-5 doi: 10.3399/bjgp16X686485

337. Care Quality Commission. *Monitoring GP practices.* 2015. https://www.cqc.org.uk/content/monitoring-gp-practices.

338. Gribben B, Coster G, Pringle M*, et al.* Quality of care indicators for population-based primary care in New Zealand. *N Z Med J* 2002;115(1151):163-6

339. Donaldson MS. *Measuring the quality of health care*: National Academies Press, 1999.

340. Zegers M, de Bruijne MC, Spreeuwenberg P*, et al.* Quality of patient record keeping: an indicator of the quality of care? *BMJ Qual Saf* 2011;20(4):314-8 doi: 10.1136/bmjqs.2009.038976

341. Haggerty JL. Ordering the chaos for patients with multimorbidity. *BMJ* 2012;345:e5915 doi: 10.1136/bmj.e5915

342. Brand CA, Ackerman IN, Bohensky MA*, et al.* Chronic disease management: a review of current performance across quality of care domains and opportunities for improving osteoarthritis care. *Rheum Dis Clin North Am* 2013;39(1):123-43 doi: 10.1016/j.rdc.2012.10.005

343. Brennan C, Harkins V, Perry IJ. Management of diabetes in primary care: a structured-care approach. *Eur J Gen Pract* 2008;14(3-4):117-22 doi: 10.1080/13814780802689154

344. Roos EM, Skou ST. Good Life with osteoArthritis in Denmark (GLA: D): Implementation of evidence-based care for knee and hip osteoarthritis into clinical practice. *Bone and Joint Decade World Summit 2015*

345. Smink AJ, van den Ende CH, Vliet Vlieland TP*, et al.* "Beating osteoARThritis": development of a stepped care strategy to optimize utilization and timing of non-surgical treatment

modalities for patients with hip or knee osteoarthritis. *Clin Rheumatol* 2011;30(12):1623-9 doi: 10.1007/s10067-011-1835-x

346. Smink AJ, Dekker J, Vliet Vlieland TP*, et al.* Health care use of patients with osteoarthritis of the hip or knee after implementation of a stepped-care strategy: an observational study. *Arthritis Care Res (Hoboken)* 2014;66(6):817-27 doi: 10.1002/acr.22222

347. Buszewicz M, Rait G, Griffin M*, et al.* Self management of arthritis in primary care: randomised controlled trial. *BMJ* 2006;333(7574):879 doi: 10.1136/bmj.38965.375718.80

348. Østerås N, van Bodegom-Vos L, Dziedzic K*, et al.* Implementing international osteoarthritis treatment guidelines in primary health care: study protocol for the SAMBA stepped wedge cluster randomized controlled trial. *Implement Sci* 2015;10:165 doi: 10.1186/s13012-015-0353-7

349. Blackburn S, Higginbottom A, Taylor R*, et al.* Patient-reported quality indicators for osteoarthritis: a patient and public generated self-report measure for primary care. *Research Involvement and Engagement* 2016;2(1):1-20 doi: 10.1186/s40900-016-0019-x

350. Gillam SJ, Siriwardena AN, Steel N. Pay-for-performance in the United Kingdom: impact of the quality and outcomes framework: a systematic review. *Ann Fam Med* 2012;10(5):461-8 doi: 10.1370/afm.1377

351. Gillam S, Steel N. The Quality and Outcomes Framework--where next? *BMJ* 2013;346:f659 doi: 10.1136/bmj.f659

352. Roland M, Campbell S. Successes and failures of pay for performance in the United Kingdom. *N Engl J Med* 2014;370(20):1944-9 doi: 10.1056/NEJMhpr1316051

353. Scarborough P, Bhatnagar P, Wickramasinghe KK*, et al.* The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006-07 NHS costs. *J Public Health (Oxf)* 2011;33(4):527-35 doi: 10.1093/pubmed/fdr033

354. National Institute for Health & Care Excellence. *NICE guidance uptake data* 2016. https://www.nice.org.uk/About/What-we-do/Into-practice/Measuring-the-uptake-of-NICE-guidance/Uptake-data.

355. Pearce-Smith N. Unwarranted Variation: a Reading List produced by QIPP Right Care. 2011. http://www.rightcare.nhs.uk/downloads/ER_unwarranted_variation_aug_2011.pdf.

356. National Clinical Guideline Centre. Osteoarthritis: Care and management in adults. 2014

357. Blanco FJ, Ruiz-Romero C. New targets for disease modifying osteoarthritis drugs: chondrogenesis and Runx1. *Ann Rheum Dis* 2013;72(5):631-4 doi: 10.1136/annrheumdis-2012-202652

358. Abraham NS, El-Serag HB, Johnson ML*, et al.* National adherence to evidence-based guidelines for the prescription of nonsteroidal anti-inflammatory drugs. *Gastroenterology* 2005;129(4):1171-8 doi: 10.1053/j.gastro.2005.08.003

359. Alghamdi MA, Olney S, Costigan P. Exercise treatment for osteoarthritis disability. *Ann Saudi Med* 2004;24(5):326-31

360. Allen KD, Oddone EZ, Stock JL*, et al.* The Self-Management of OsteoArthritis in Veterans (SeMOA) Study: design and methodology. *Contemp Clin Trials* 2008;29(4):596-607 doi: 10.1016/j.cct.2007.11.004

361. American College of Rheumatology. *Quality Measures* 2011. http://www.rheumatology.org/practice/clinical/quality/quality.asp.

362. Arthritis and Musculoskeletal Alliance. *Standards of Care for people with Osteoarthritis*, 2004.

363. Askari M, Wierenga PC, Eslami S*, et al.* Assessing quality of care of elderly patients using the ACOVE quality indicator set: a systematic review. *PLoS One* 2011;6(12):e28631 doi: 10.1371/journal.pone.0028631

364. Barlow J. How to use education as an intervention in osteoarthritis. *Best Pract Res Clin Rheumatol* 2001;15(4):545-58 doi: 10.1053/berh.2001.0172

365. Baumann M, Euller-Ziegler L, Guillemin F. Evaluation of the expectations osteoarthritis patients have concerning healthcare, and their implications for practitioners. *Clinical and Experimental Rheumatology* 2007;25(3):404-09

366. Bierma-Zeinstra SM, Lipschart S, Njoo KH*, et al.* How do general practitioners manage hip problems in adults? *Scand J Prim Health Care* 2000;18(3):159-64

367. Bijlsma JW. Analgesia and the patient with osteoarthritis. *Am J Ther* 2002;9(3):189-97

368. Blanda MP. Pharmacologic issues in geriatric emergency medicine. *Emerg Med Clin North Am* 2006;24(2):449-65, viii doi: 10.1016/j.emc.2006.01.007

369. Bliddal H, Christensen R. The management of osteoarthritis in the obese patient: practical considerations and guidelines for therapy. *Obes Rev* 2006;7(4):323-31 doi: 10.1111/j.1467-789X.2006.00252.x

370. Bloodworth D. Issues in opioid management. *Am J Phys Med Rehabil* 2005;84(3 Suppl):S42-55

371. Boyd CM, Darer J, Boult C*, et al.* Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 2005;294(6):716-24 doi: 10.1001/jama.294.6.716

372. Brand C. Translating evidence into practice for people with osteoarthritis of the hip and knee. *Clin Rheumatol* 2007;26(9):1411-20 doi: 10.1007/s10067-007-0633-y

373. Brand CA. The role of self-management in designing care for people with osteoarthritis of the hip and knee. *Med J Aust* 2008;189(10 Suppl):S25-8

374. Brand C, Cox S. Systems for implementing best practice for a chronic disease: management of osteoarthritis of the hip and knee. *Intern Med J* 2006;36(3):170-9 doi: 10.1111/j.1445-5994.2006.01018.x

375. Brooks P, Conaghan PG. Principles of managing musculoskeletal conditions. Preface. *Best Pract Res Clin Rheumatol* 2008;22(3):385-6 doi: 10.1016/j.berh.2008.03.001

376. Brosseau L, Wells GA, Tugwell P*, et al.* Ottawa panel evidence-based clinical practice guidelines for therapeutic exercises and manual therapy in the management of osteoarthritis. *Phys Ther* 2005;85(9):907-71

377. Cadogan MP, Schnelle JF, Yamamoto-Mitani N*, et al.* A Minimum Data Set Prevalence of Pain Quality Indicator: Is It Accurate and Does It Reflect Differences in Care Processes? *The*

Journals of Gerontology Series A: Biological Sciences and Medical Sciences 2004;59(3):M281-M85 doi: 10.1093/gerona/59.3.M281

378. Chassany O, Boureau F, Liard F*, et al.* Effects of training on general practitioners' management of pain in osteoarthritis: a randomized multicenter study. *J Rheumatol* 2006;33(9):1827-34

379. Cieza A, Ewert T, Ustun TB*, et al.* Development of ICF Core Sets for patients with chronic conditions. *J Rehabil Med* 2004(44 Suppl):9-11 doi: 10.1080/16501960410015353

380. Conaghan PG, Brooks P. How to manage musculoskeletal conditions. *Best Pract Res Clin Rheumatol* 2008;22(3):387-94 doi: 10.1016/j.berh.2008.02.004

381. Cornali C, Franzoni S, Gatti S*, et al.* Diagnosis of chronic pain caused by osteoarthritis and prescription of analgesics in patients with cognitive impairment. *J Am Med Dir Assoc* 2006;7(1):1-5 doi: 10.1016/j.jamda.2005.06.010

382. Cox ER, Motheral B, Frisse M*, et al.* Prescribing COX-2s for patients new to cyclo-oxygenase inhibition therapy. *Am J Manag Care* 2003;9(11):735-42

383. Curtis JR, Saag KG. Evaluating and improving the quality of care in rheumatic disease. *Expert Rev Pharmacoecon Outcomes Res* 2004;4(4):429-39 doi: 10.1586/14737167.4.4.429

384. Doherty M, Dougados M. Evidence-based management of osteoarthritis: practical issues relating to the data. *Best Pract Res Clin Rheumatol* 2001;15(4):517-25 doi: 10.1053/berh.2001.0170

385. Dominick KL, Baker TA. Racial and ethnic differences in osteoarthritis: prevalence, outcomes, and medical care. *Ethn Dis* 2004;14(4):558-66

386. Dominick KL, Bosworth HB, Dudley TK*, et al.* Patterns of opioid analgesic prescription among patients with osteoarthritis. *J Pain Palliat Care Pharmacother* 2004;18(1):31-46

387. Dooley P, Martin R. Corticosteroid injections and arthrocentesis. *Can Fam Physician* 2002;48:285-92

388. Duran-Barragan S, Russell AS. Use of nonsteroidal anti-inflammatory drugs and coxibs in the elderly. Are we following the guidelines? *Clinical Rheumatology* 2008;27(9):1081

389. Emejuaiwe N, Jones AC, Ibrahim SA*, et al.* Disparities in joint replacement utilization: a quality of care issue. *Clin Exp Rheumatol* 2007;25(6 Suppl 47):44-9

390. Escobar A, Quintana JM, Arostegui I*, et al.* Development of explicit criteria for total knee replacement. *Int J Technol Assess Health Care* 2003;19(1):57-70

391. Feldt KS. The complexity of managing pain for frail elders. *J Am Geriatr Soc* 2004;52(5):840-1 doi: 10.1111/j.1532-5415.2004.52228.x

392. Ganz DA, Wenger NS, Roth CP*, et al.* The effect of a quality improvement initiative on the quality of other aspects of health care: the law of unintended consequences? *Med Care* 2007;45(1):8-18 doi: 10.1097/01.mlr.0000241115.31531.15

393. Glazier RH, Badley EM, Wright JG*, et al.* Patient and provider factors related to comprehensive arthritis care in a community setting in Ontario, Canada. *J Rheumatol* 2003;30(8):1846-50

394. Gogovor A, Savoie M, Moride Y*, et al.* Contemporary disease management in Quebec. *Healthc Q* 2008;11(1):30-7

395. Gordon J, Rockwood K, Powell C. Assessing patients' views of clinical changes. *JAMA* 2000;283(14):1824-5

396. Gorevic PD. Osteoarthritis. A review of musculoskeletal aging and treatment issues in geriatric patients. *Geriatrics* 2004;59(8):28-32; quiz 35

397. Gossec L, Hawker G, Davis AM*, et al.* OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. *J Rheumatol* 2007;34(6):1432-5

398. Graudins LV, Gazarian M. Celecoxib use and overuse - Too much 'celebration'? *Journal of Pharmacy Practice and Research* 2002;32(2):104

399. Hall AM. A nurse-led service to provide information on osteoarthritis. *Nurs Times* 2005;101(3):38-9

400. Higashi T, Wenger NS, Adams JL*, et al.* Relationship between number of medical conditions and quality of care. *N Engl J Med* 2007;356(24):2496-504 doi: 10.1056/NEJMsa066253

401. Holden MA, Nicholls EE, Hay EM*, et al.* Physical therapists' use of therapeutic exercise for patients with clinical knee osteoarthritis in the United kingdom: in line with current recommendations? *Phys Ther* 2008;88(10):1109

402. Hollenack KA, Cranmer KW, Zarowitz BJ*, et al.* The application of evidence-based principles of care in older persons (issue 4): pain management. *J Am Med Dir Assoc* 2007;8(3 Suppl 2):e77-85 doi: 10.1016/j.jamda.2006.12.016

403. Holman H, Bloch DA, Lorig K*, et al.* Assessing patients' views of clinical changes - Reply. *JAMA* 2000;283(14):1825-25

404. Hunter DJ. Quality of osteoarthritis care for community-dwelling older adults. *Clin Geriatr Med* 2010;26(3):401-17 doi: 10.1016/j.cger.2010.03.003

405. Hunter DJ, Neogi T, Hochberg MC. Quality of osteoarthritis management and the need for reform in the US. *Arthritis Care Res (Hoboken)* 2011;63(1):31-8 doi: 10.1002/acr.20278

406. Jansen MJ, Hendriks EJ, Oostendorp RA*, et al.* Quality indicators indicate good adherence to the clinical practice guideline on "Osteoarthritis of the hip and knee" and few prognostic factors influence outcome indicators: a prospective cohort study. *Eur J Phys Rehabil Med* 2010;46(3):337-45

407. Jawad AS. Analgesics and osteoarthritis: are treatment guidelines reflected in clinical practice? *Am J Ther* 2005;12(1):98-103

408. Jones K, Piterman L. Promoting best practice in general practitioner management of osteoarthritis of the hip and knee: Arthritis and musculoskeletal quality improvement program (AMQuIP). *Aust J Prim Health* 2007;13(2):104-12 doi: 10.1071/PY07028

409. Jordan KM, Sawyer S, Coakley P*, et al.* The use of conventional and complementary treatments for knee osteoarthritis in the community. *Rheumatology (Oxford)* 2004;43(3):381-4 doi: 10.1093/rheumatology/keh045

410. Kane RL, Saleh KJ, Wilt TJ*, et al.* Total knee replacement. *Evid Rep Technol Assess (Summ)* 2003(86):1-8

411. Kartal M, Maral I, Coskun O. Prescribing pattern of general practitioners for osteoarthritis in primary care settings in Bolu, Turkey. *Saudi Med J* 2007;28(12):1885-9

412. Kirkness CS, Yu J, Asche CV. The effect on comorbidity and pain in patients with osteoarthritis. *J Pain Palliat Care Pharmacother* 2008;22(4):336-48 doi: 10.1080/15360280802536649

413. Koehn CL, Esdaile JM. Patient education and self-management of musculoskeletal diseases. *Best Pract Res Clin Rheumatol* 2008;22(3):395-405 doi: 10.1016/j.berh.2008.02.006

414. Koyama Y, Miyashita M, Irie S*, et al.* A study of disease management activities of hip osteoarthritis patients under conservative treatment. *Journal of Orthopaedic Nursing* 2008;12(2):75

415. Krishnan LL, Suarez-Almazor ME. Evidence-based rheumatology practice. *Curr Opin Rheumatol* 2005;17(2):117-23

416. Labelle M, Beaulieu M, Paquette D*, et al.* An integrated approach to improving appropriate use of anti-inflammatory medication in the treatment of osteoarthritis in Quebec (Canada): the CURATA model. *Med Teach* 2004;26(5):463-70 doi: 10.1080/0142159042000218669

417. Lai HY, Chen YC, Chen TJ*, et al.* Intra-articular hyaluronic acid for treatment of osteoarthritis: a nationwide study among the older population of Taiwan. *BMC Health Serv Res* 2008;8:24 doi: 10.1186/1472-6963-8-24

418. Laine L, Connors L, Griffin MR*, et al.* Prescription rates of protective co-therapy for NSAID users at high GI risk and results of attempts to improve adherence to guidelines. *Aliment Pharmacol Ther* 2009;30(7):767-74 doi: 10.1111/j.1365-2036.2009.04090.x

419. Landsberg PG, Pillans PI, Radford JM. Evaluation of cyclooxygenase 2 inhibitor use in patients admitted to a large teaching hospital. *Internal Medicine Journal* 2003;33(5-6):225-28 doi: 10.1046/j.1445-5994.2003.00392.x

420. Lanza FL, Chan FK, Quigley EM*, et al.* Guidelines for prevention of NSAID-related ulcer complications. *Am J Gastroenterol* 2009;104(3):728-38 doi: 10.1038/ajg.2009.115

421. Lapane KL, Spooner JJ, Pettitt D. The effect of nonsteroidal anti-inflammatory drugs on the use of gastroprotective medication in people with arthritis. *Am J Manag Care* 2001;7(4):402-8

422. Lee YC, Katz JN. Shared decision making for total joint replacement: the physician's role. *Journal of Musculoskeletal Medicine* 2007;24(8):335

423. LeLorier J, Fitzsimon C, Keresteci M*, et al.* Drug utilization review of celecoxib in Ontario. *Rheumatology (Oxford)* 2003;42 Suppl 3:iii11-6 doi: 10.1093/rheumatology/keg493

424. Li LC, Koehn C, Lehman AJ. People with arthritis and their families in rehabilitation, care and research. *Curr Opin Rheumatol* 2007;19(2):174-8 doi: 10.1097/BOR.0b013e3280327933

425. LoBuono C. AHRQ issues two guides on osteoarthritis. *Drug Topics* 2007;151(5)

426. Maillefert JF, Roy C, Cadet C*, et al.* Factors influencing surgeons' decisions in the indication for total joint replacement in hip osteoarthritis in real life. *Arthritis Rheum* 2008;59(2):255-62 doi: 10.1002/art.23331

427. March L, Amatya B, Osborne RH*, et al.* Developing a minimum standard of care for treating people with osteoarthritis of the hip and knee. *Best Pract Res Clin Rheumatol* 2010;24(1):121-45 doi: 10.1016/j.berh.2009.10.002

428. Marcum ZA, Perera S, Donohue JM*, et al.* Analgesic Use for Knee and Hip Osteoarthritis in Community-Dwelling Elders. *Pain Med* 2011;12(11):1628-36 doi: 10.1111/j.1526-4637.2011.01249.x

429. Marks R, Allegrante JP. Chronic osteoarthritis and adherence to exercise: a review of the literature. *J Aging Phys Act* 2005;13(4):434-60

430. Mazieres B, Scmidely N, Hauselmann HJ*, et al.* Level of acceptability of EULAR recommendations for the management of knee osteoarthritis by practitioners in different European countries. *Ann Rheum Dis* 2005;64(8):1158-64 doi: 10.1136/ard.2003.009431

431. McHugh GA, Silman AJ, Luker KA. Quality of care for people with osteoarthritis: a qualitative study. *J Clin Nurs* 2007;16(7B):168-76 doi: 10.1111/j.1365-2702.2007.01885.x

432. McHugh GA, Campbell M, Luker KA. Quality of care for individuals with osteoarthritis: a longitudinal study. *J Eval Clin Pract* 2012;18(3):534-41 doi: 10.1111/j.1365-2753.2010.01616.x

433. McPherson KM, Reid DA. New roles in health care: What are the key questions? *Med J Aust* 2007;186(12):614

434. Merkle D, McDonald DD. Use of recommended osteoarthritis pain treatment by older adults. *J Adv Nurs* 2009;65(4):828-35 doi: 10.1111/j.1365-2648.2008.04940.x

435. Min LC, Reuben DB, MacLean CH*, et al.* Predictors of overall quality of care provided to vulnerable older people. *J Am Geriatr Soc* 2005;53(10):1705-11 doi: 10.1111/j.1532-5415.2005.53520.x

436. Mitchell HL, Hurley MV. Management of chronic knee pain: a survey of patient preferences and treatment received. *BMC Musculoskelet Disord* 2008;9:123 doi: 10.1186/1471-2474-9-123

437. Morewitz SJ. *Chronic diseases and health care: New trends in diabetes, arthritis, osteoporosis, fibromyalgia, low back pain, cardiovascular disease and cancer*, 2006.

438. Newman ED, Harrington JT. Redesigning the care of rheumatic diseases at the practice and system levels - Part 2: System level process improvement (Redesign 201). *Clinical and Experimental Rheumatology* 2007;25(6):S64-S68

439. Oliver S. Multidisciplinary disease management in rheumatology. *Prof Nurse* 2003;19(3):137-41

440. O'Reilly S, Doherty M. Lifestyle changes in the management of osteoarthritis*. Best Pract Res Clin Rheumatol* 2001;15(4):559-68 doi: 10.1053/berh.2001.0173

441. Østerås N, Hagen KB, Garratt A*, et al.* Quality of osteoarthritis care: test-retest reliability and feasibility of the osteoarthritis quality indicator questionnaire. *Osteoarthritis Cartilage* 2012;20:S185-S86 doi: 10.1016/j.joca.2012.02.287

442. Osterhaus JT, Dedhiya SD, Ernst ME*, et al.* Health outcomes assessment in community pharmacy practices: a feasibility project. *Arthritis Rheum* 2002;47(2):124-31

443. Pendleton A, Arden N, Dougados M*, et al.* EULAR recommendations for the management of knee osteoarthritis: report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Ann Rheum Dis* 2000;59(12):936-44

444. Peter WF, van der Wees P, Verhoef J*, et al.* ACR Meeting. *Arthritis Rheum* 2011;63(S10):1-1 doi: 10.1002/art.33310

445. Pitt VJ, O'Connor D, Green S. Referral of people with osteoarthritis to self-management programmes: barriers and enablers identified by general practitioners. *Disabil Rehabil* 2008;30(25):1938 - 46 doi: 10.1080/09638280701774233

446. Porcheret M, Jordan K, Croft P*, et al.* Treatment of knee pain in older adults in primary care: development of an evidence-based model of care. *Rheumatology (Oxford)* 2007;46(4):638-48 doi: 10.1093/rheumatology/kel340

447. Price-Forbes AN, Callaghan R, Allen ME*, et al.* A regional audit of the use of COX-2 selective non-steroidal anti-inflammatory drugs (NSAIDs) in rheumatology clinics in the West Midlands, in relation to NICE guidelines. *Rheumatology (Oxford)* 2005;44(7):921-4 doi: 10.1093/rheumatology/keh642

448. Quintana JM, Arostegui I, Azkarate J*, et al.* Evaluation of explicit criteria for total hip joint replacement. *J Clin Epidemiol* 2000;53(12):1200-08 doi: 10.1016/S0895-4356(00)00244-4

449. Quintana JM, Arostegui I, Azkarate J*, et al.* Evaluation by explicit criteria of the use of total hip joint replacement. *Rheumatology (Oxford)* 2000;39(11):1234-41 doi: 10.1093/rheumatology/39.11.1234

450. Quintana JM, Escobar A, Azkarate J*, et al.* Appropriateness of total hip joint replacement. *Int J Qual Health Care* 2005;17(4):315-21 doi: 10.1093/intqhc/mzi047

451. Quintana JM, Bilbao A, Escobar A*, et al.* Decision trees for indication of total hip replacement on patients with osteoarthritis. *Rheumatology (Oxford)* 2009;48(11):1402-9 doi: 10.1093/rheumatology/kep264

452. Reeves D, Campbell SM, Adams J*, et al.* Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. *Med Care* 2007;45(6):489-96 doi: 10.1097/MLR.0b013e31803bb479

453. Rosemann T, Laux G, Szecsenyi J*, et al.* The Chronic Care Model: congruency and predictors among primary care patients with osteoarthritis. *Qual Saf Health Care* 2008;17(6):442-6 doi: 10.1136/qshc.2007.022822

454. Saag KG, Olivieri JJ, Patino F*, et al.* Measuring quality in arthritis care: the Arthritis Foundation's quality indicator set for analgesics. *Arthritis Rheum* 2004;51(3):337-49 doi: 10.1002/art.20422

455. Saliba D, Solomon D, Rubenstein L*, et al.* Feasibility of quality indicators for the management of geriatric syndromes in nursing home residents. *J Am Med Dir Assoc* 2005;6(3 Suppl):S50-9 doi: 10.1016/j.jamda.2005.03.023

456. Schilling ML. Pain management in older adults. *Curr Psychiatry Rep* 2003;5(1):55-61

457. Singh G, Heller MD, Silver D*, et al.* Current approach to safe osteoarthritis pharmacotherapy. *Consultant* 2003;43(8):3

458. Solomon DH, Schaffer JL, Katz JN*, et al.* Can history and physical examination be used as markers of quality? An analysis of the initial visit note in musculoskeletal care. *Med Care* 2000;38(4):383-91

459. Strand V. Measuring quality in arthritis care: the Arthritis Foundation's quality indicator set for rheumatoid arthritis. *Semin Arthritis Rheum* 2006;35(4):205-7 doi: 10.1016/j.semarthrit.2005.09.006

460. Sturkenboom MC, Burke TA, Dieleman JP*, et al.* Underutilization of preventive strategies in patients receiving NSAIDs. *Rheumatology (Oxford)* 2003;42 Suppl 3:iii23-31 doi: 10.1093/rheumatology/keg495

461. Vandenberghe H, Van Casteren V, Jonckheer P*, et al.* Quality of care assessment using GPs' electronic patient records: do we need data from home visits? *Stud Health Technol Inform* 2004;110:35-41

462. Vandenberghe HE, Van Casteren V, Jonckheer P*, et al.* Collecting information on the quality of prescribing in primary care using semi-automatic data extraction from GPs' electronic medical records. *Int J Med Inform* 2005;74(5):367-76 doi: 10.1016/j.ijmedinf.2005.02.004

463. Webster JR. How and why we should move our geriatric cheese. *Ann Intern Med* 2001;135(8):641-42

464. Wierenga PC, Klopotowska JE, Smorenburg SM*, et al.* Quality indicators for in-hospital pharmaceutical care of Dutch elderly patients: development and validation of an ACOVE-based quality indicator set. *Drugs Aging* 2011;28(4):295-304 doi: 10.2165/11587700-000000000-00000

465. Wilcock M, Davis M, Sampson B. The prescribing of non-steroidal anti-inflammatory drugs in general practice can be improved. *Pharmacy in Practice* 2004;14(6):164-68

466. Yazdany J, MacLean CH. Quality of care in the rheumatic diseases: current status and future directions. *Curr Opin Rheumatol* 2008;20(2):159-66 doi: 10.1097/BOR.0b013e3282f50ec4

467. Zingmond DS, Wilber KH, Maclean CH*, et al.* Measuring the quality of care provided to community dwelling vulnerable elders dually enrolled in Medicare and Medicaid. *Med Care* 2007;45(10):931-8 doi: 10.1097/MLR.0b013e318127143e

Appendices

Appendix A. Peer-reviewed publication resulting from the systematic review of quality indicators for the primary care of OA

OPEN ACCESS

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 Dziedzic[1]

Handling editor Tore K Kvien

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2013-203913).

[1]Arthritis Research UK Primary Care Centre, Research Institute for Primary Care & Health Sciences, Keele University, Keele, Staffordshire, UK
[2]Earnswood Medical Centre, Eagle Bridge Health & Well Being Centre, Crewe, Cheshire, UK

Correspondence to
Dr J J Edwards, Arthritis Research UK Primary Care Centre, Research Institute for Primary Care & Health Sciences, Keele University, Keele, Staffordshire ST5 5BG, UK;
j.j.edwards@keele.ac.uk

Received 8 May 2013
Revised 9 September 2013
Accepted 6 November 2013
Published Online First 28 November 2013

To cite: Edwards JJ, Khanna M, Jordan KP, *et al*. *Ann Rheum Dis* 2015;**74**:490–498.

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 Hochberg[18] and Strömbeck *et al*[19] 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 al*[20] 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

Table 1 MEDLINE search strategy

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

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.

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 piloted 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 homes[42] and home-based primary care[45]) 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 2 Indicator studies

Indicator set Author and date	Truth		Target population	Proposed method of measurement	Testing or implementation	Reliability	Feasibility
	Evidence synthesis	Consensus method					
RAND Quality of Care Assessment Tools (RAND QA) Moore (2000)[29]	Literature review, not specified to be systematic	RAND Appropriateness Method	Not specified	Medical record review	McGlynn *et al*[30] Asch *et al*[31]	Tested in McGlynn *et al*[30] and Asch *et al*[31] —in a 4% sample reabstraction reliability was substantial at 3 levels: presence of a condition (κ=0.83), indicator eligibility (κ=0.76) and indicator scoring κ=0.80)	McGlynn *et al*[30] 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 al*[31] 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)[32 33]	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 al*[34] Chodosh *et al*[35] Higashi *et al*[36 37] Ganz *et al*[38] MacLean *et al*[39] Østerås *et al*[40]	Tested in Wenger *et al*[34]—10% sample reabstraction: overall error rate was 1.6%; also in Chodosh *et al*[35]—'Inter-rater reliability of chart abstraction for eligibility and scoring of indicators was 95%.'; Higashi *et al*[36 37]—10% reabstraction sample showed 97% identical eligibility and 95% identical eligibility and quality score; MacLean *et al*[39]—10% reabstraction sample with κ=0.85 (93% agreement); Østerås *et al*[40] questionnaire-based test-retest κ=0.20–0.80, % exact agreement from 62–90%	Wenger *et al*[34] implemented the indicators in community dwelling VEs in the USA—medical record abstraction by trained nurses supplemented for some indicators by telephone interview. Chodosh *et al*,[35] Higashi *et al*[36 37] and MacLean *et al*[39] used the same population and methods. Ganz *et al* used a similar population and methodology at a different time point[38] Østerås *et al*[40] 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	Steel *et al*,[12 14] Broadbent *et al*,[13] Østerås *et al*[40]	Tested in Steel *et al*[12]—(κ=0.8, 95% CI=0.7 to 0.9); Østerås *et al*[40] questionnaire-based test-retest κ=0.20–0.80, % exact agreement from 62–90%	Broadbent *et al*[13] and Steel *et al*[12] separately implemented indicators in UK general practice, using medical record review (computerised and paper notes). Østerås *et al*[40] 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 al*[43] Zingmond *et al*[44]	Tested in Cadogan *et al*[43]—κ=0.65–1.00 and percentage agreement 80–100 where κ could not be calculated (numbers too low)	Cadogan *et al*[43] implemented indicators in 30 nursing homes in California using medical record review. Zingmond *et al*[44] implemented using Medicare and Medicaid eligibility and claims data and a nursing home minimum dataset
ACOVE-1 adapted for the HPCQI 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	No feasibility testing identified
ACOVE-3 ACOVE investigators (2007)[46–48]	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 al*[40]	Østerås *et al*[40] questionnaire-based test-retest κ=0.20–0.80, % exact agreement from 62–90%	Østerås *et al*[40] implemented some of these indicators in a modified patient self-report format

Continued

Table 2 Continued

Indicator set Author and date	Truth: Evidence synthesis	Truth: Consensus method	Target population	Proposed method of measurement	Testing or implementation	Reliability	Feasibility
	a project member knowledgeable about systematic reviews and indicator development		death or functional decline over a 2-year period				
QIGP Underwood (2002)[49]	Various sources used (Cochrane, DARE, Medline) but not clear how the evidence was assembled. Cites meta-analyses, systematic reviews, randomised controlled trials	Not stated	Not specified	Not specified	Kirk *et al*[50] Steel *et al*[12] Broadbent *et al*[13] Østerås *et al*[40]	Tested for non-OA indicators in Kirk *et al*[50]—as OA was not included in this exercise, it is not known what degree of reliability exists for these indicators; Østerås *et al*[40] questionnaire-based test-retest κ=0.20–0.80, % exact agreement from 62–90%	Kirk *et al*[50] implemented in 16 UK general practices in two areas using data from electronic and paper records. Steel *et al*[12] and Broadbent *et al*[13] separately implemented the NSAID indicator in UK general practice, using medical record review (computerised and paper notes). Østerås *et al*[40] implemented some of these indicators in a modified patient self-report format
Arthritis Foundation Arthritis Foundation 2004[51 52]	Comprehensive literature search and expert opinion	Modified RAND/ UCLA Appropriateness Method	Patients with OA	Not specified	Li *et al*[53]	No reliability testing identified	Li *et al*[53] used a postal survey in Canada (sampling frame from an administrative database in British Columbia) to assess non-pharmacological indicators
PCPI (2006)[54]	PCPI website refers to a methodology committee but no specific information in the indicator set to identify how it was developed		All patients aged ≥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	No feasibility testing identified
EUMUSC.net (2012)[55]	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	No feasibility 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.

Table 3 Narrative synthesis of exemplar indicators and their feasibility for use in primary care

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)[46–48]	RAND QA,[29] ACOVE-1,[32 33] and as adapted (ELSA,[41] HPCQI[45]), Arthritis Foundation,[51 52] PCPI,[54] EUMUSC.net[55]	12 13 30 31 34 35 38 40 Requires change in routine coding to improve capture of this information
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)[46–48]	RAND QA,[29] ACOVE-1,[32 33] and as adapted (ELSA,[41] HPCQI[45]), Arthritis Foundation,[51 52] PCPI[54]	12 13 30 31 35 38 40 Requires change in routine coding to improve capture of this information
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)[51 52]	ACOVE-1 (2 variations—new and pre-existing disease),[32 33] and as adapted (ELSA[41]), EUMUSC.net[55]	12–14 34 38 40 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)[46–48]	*Initial recommendation* RAND QA,[29] ACOVE-1 (indicators for new and pre-existing disease),[32 33] and as adapted (ELSA,[41] ACOVE/NH,[42] HPCQI[45]), Arthritis Foundation,[51 52] PCPI,[54] EUMUSC.net[55] *Annual review* RAND QA,[29] ACOVE-1[32 33]	*Initial recommendation* 14 30 31 34 37 38 40 43 53 *Annual review* 53 Requires change in routine coding to improve capture of this information
Weight loss 1 (ACR (hip, knee), NICE, OARSI	IF a VE is obese (body mass index (BMI) ≥30 kg/m^2), THEN he or she should be advised annually to lose weight… (ACOVE-3)[46–48]	Arthritis Foundation[51 52]	No implementation studies identified for this indicator. Should be captured from existing weight and health promotion records
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/m^2), 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)[51 52]	EUMUSC.net[55]	40 53 Consider a lower BMI threshold of 25 kg/m^2 for consistency with the usual definition of 'overweight'. Should be captured from existing weight and health promotion records.
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)[46–48]	Arthritis Foundation,[51 52] EUMUSC.net[55]	40 53 Requires change in routine coding to improve capture of this information
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)[46–48]	Arthritis Foundation,[51 52] EUMUSC.net[55]	40 53 Requires change in routine coding to improve capture of this information
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)[46–48]	RAND QA,[29] ACOVE-1,[32 33] and as adapted (ELSA,[41] ACOVE/NH,[42] HPCQI[45]), QIGP,[49] Arthritis Foundation,[51 52]	12–14 30 31 34 36–38 40 43 50 Requires change in routine coding to capture over-the-counter drug use
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)[51 52]	ACOVE-1,[32 33] and as adapted (ELSA,[41] ACOVE/NH,[42] HPCQI[45])	12 13 34 36 43 Requires change in routine coding to capture over-the-counter drug use
Oral NSAIDs 1 (all guidance)	If NSAIDs are considered, ibuprofen should be considered for first-line treatment unless contraindicated or intolerant.* (QIGP)[49]	Modifications exist in implementation studies: Steel *et al*,[12] Broadbent *et al*[13] to include use of COX-2 selective drugs	12 13 50 Requires change in routine coding to capture over-the-counter drug use.
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)[54]	Two indicators from ACOVE-3 refer to risks from NSAIDs and aspirin to be 'discussed and documented',[46–48] EUMUSC.net[55]	12 13 Requires change in routine coding to capture over-the-counter drug use

Continued

Table 3 Continued

Overarching theme (source)	'Exemplar' indicator	Reproducibility (other sources of similar indicators)	Implementation references and comment on feasibility
Gastroprotection (EULAR (all sites), NICE, OARSI)	IF a VE with a risk factor for GI bleeding (aged ≥75, peptic ulcer disease, history of GI bleeding, warfarin use, chronic glucocorticoid use) is treated with a non-selective NSAID, THEN he or she should be treated concomitantly with misoprostol or a PPI. (ACOVE-3)[46–48]	ACOVE-1,[32 33] ACOVE-3 [46–48] (NSAIDs, and aspirin), QIGP,[49] PCPI[54]	[34–39 44 50] Should be captured from existing electronic prescribing records
Specialist referral (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, BECAUSE joint surgery may reduce pain and improve functional status and quality of life. (ACOVE-3)[46–48]	RAND QA,[29] ACOVE-1,[32 33] and as adapted (ELSA[41]), Arthritis Foundation,[51 52] QIGP,[49] EUMUSC.net[55]	[12–14 37 38 40 50] It would be feasible to capture the presence of non-surgical therapy indicators in the record, though routine data sources cannot be used to determine the need for a surgical opinion reliably

*It should be noted that different sources of guidance offer varying recommendations about the use of specific NSAIDs; in the UK, NICE recommend a standard NSAID or COX-2 inhibitor (other than etoricoxib 60mg) to be coprescribed with a PPI.[8]
ACOVE, Assessing Care of Vulnerable Elders; ACR, American College of Rheumatology; ADL, activities of daily living; COX, cyclooxygenase, ELSA, English Longitudinal Study of Ageing; EULAR, European League Against Rheumatism; EUMUSC.net, European Musculoskeletal Conditions Surveillance and Information Network; GI, gastrointestinal; HPCQI, Home-based Primary Care Quality Initiative; NH, nursing home; NICE, National Institute for Health and Care Excellence; NSAIDs, non-steroidal anti-inflammatories; OA, Osteoarthritis; OARSI, Osteoarthritis Research Society International; OTC, over the counter; PCPI, Physician Consortium for Performance Improvement; PPI, proton-pump inhibitor; QA, Quality Assessment; QIGP, Quality Indicators for General Practice; VE, vulnerable elder.

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 specified 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?"[38]; 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.

The EUMUSC.net team includes an education indicator aimed at clinicians, which we did not include as it is not a patient-focused 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[55]; one self-report indicator implemented also included patients with hand, hip or knee OA[40]; 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[43]; 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,

Table 4 Proposed indicators for primary care implementation

Overarching theme	Proposal for primary care implementation
Holistic Assessment: Pain	% patients with a working diagnosis of OA with evidence of pain assessment within the previous 12 months
Holistic Assessment: Function	% patients with a working diagnosis of OA with evidence of function assessment within the previous 12 months
Education	% patients with a working diagnosis of OA with evidence of education or advice since diagnosis
Exercise 1	% patients with a working diagnosis of OA in the hip or knee with evidence of exercise advice or physiotherapy referral since diagnosis
Exercise 2	% patients with a working diagnosis of OA with evidence of an activity review within the previous 12 months
Weight loss 1	% patients with a BMI $\geq$30 kg/m^2 who have a record of weight loss advice within the previous 12 months
Weight loss 2	% patients with a working diagnosis of OA with a BMI $\geq$25 kg/m^2 who have a record of weight loss advice within the previous 12 months
Aids and devices 1	% patients with a working diagnosis of OA with evidence of functional impairment who are recorded as receiving a referral or assessment for ambulatory assistive devices within the previous 12 months
Aids and devices 2	% patients with a working diagnosis of OA with evidence of functional impairment who are recorded as receiving a referral or assessment for assistive devices within the previous 12 months
Paracetamol 1	% patients with a working diagnosis of OA with evidence of paracetamol as the first oral analgesic prescribed or advised since diagnosis
Paracetamol 2	% patients with a working diagnosis of OA taking oral analgesics or NSAIDs with evidence that a suitable maximal dose of paracetamol was tried beforehand
Oral NSAIDs 1	% patients with a working diagnosis of OA with evidence of a standard NSAID or COX-2 inhibitor as the first oral NSAID prescribed or advised since diagnosis
Oral NSAIDs 2	% patients with a working diagnosis of OA taking an oral NSAID with a documented risk assessment prior to first prescription
Gastroprotection	% patients with a working diagnosis of OA taking an oral NSAID who are also prescribed a PPI or alternative gastroprotective agent
Specialist assessment	% patients with a record of achievement of all other applicable indicators prior to specialist referral*

*That is, the other 14 indicators above, depending on applicability of weight and therapy indicators to individual patients.
BMI, body mass index; COX, cyclooxygenase; NSAIDs, non-steroidal anti-inflammatories; OA, Osteoarthritis; PPI, proton-pump inhibitor.

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*[53] used entry to a weight-loss programme or dietetics appointment as criteria for indicator achievement and the weight loss advice self-report indicator in Østerås *et al*[40] 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*,[53] 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*.[40] 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[54] (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[54] and an implemented indicator regarding stronger analgesics (Østerås *et al*[40]) 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[54] 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.[38] Another used a patient self-report to identify failure of conservative treatment leading to referral.[40] 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.[40] 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[55]:

- 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.[56]

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.[57] 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.[58]

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 supports, 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.[8]

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 Østerås *et al*[40] 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 Olalekan Uthman, 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 JLJ 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.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 RCGP Birmingham Research Unit. Weekly Returns Service Annual Prevalence Report 2007. http://www.rcgp.org.uk/clinical-and-research/~/media/Files/CIRC/CIRC%2076%20-%2080/BRU_Annual_prevalence_report_2007.ashx
2 National Collaborating Centre for Chronic Conditions. *Osteoarthritis: national clinical guideline for care and management in adults*. London: Royal College of Physicians, 2008.
3 Hochberg MC, Altman RD, April KT, *et al*. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res* 2012;64:465–74.
4 Jordan KM, Arden NK, Doherty M, *et al*. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Ann Rheum Dis* 2003;62:1145–55.
5 Zhang W, Doherty M, Arden N, *et al*. EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2005;64:669–81.
6 Zhang W, Doherty M, Leeb BF, *et al*. EULAR evidence based recommendations for the management of hand osteoarthritis: report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007;66:377–88.
7 Zhang W, Moskowitz RW, Nuki G, *et al*. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarth Cartilage* 2008;16:137–62.
8 National Institute for Health & Clinical Excellence. *Osteoarthritis: the care and management of osteoarthritis in adults*. London: National Institute for Health & Clinical Excellence, 2008.
9 Croft P, Porcheret M, Peat G. Managing osteoarthritis in primary care: the GP as public health physician and surgical gatekeeper. *Br J Gen Pract* 2011;61:485–6.
10 Porcheret M, Jordan K, Jinks C, *et al*. Primary care treatment of knee pain—a survey in older adults. *Rheumatology* 2007;46:1694–700.
11 McHugh GA, Luker KA, Campbell M, *et al*. A longitudinal study exploring pain control, treatment and service provision for individuals with end-stage lower limb osteoarthritis. *Rheumatology* 2007;46:631–7.
12 Steel N, Maisey S, Clark A, *et al*. Quality of clinical primary care and targeted incentive payments: an observational study. *Br J Gen Pract* 2007;57:449–54.
13 Broadbent J, Maisey S, Holland R, *et al*. Recorded quality of primary care for osteoarthritis: an observational study. *Br J Gen Pract* 2008;58:839.
14 Steel N, Bachmann M, Maisey S, *et al*. Self reported receipt of care consistent with 32 quality indicators: national population survey of adults aged 50 or more in England. *BMJ* 2008;337:a957. doi:10.1136/bmj.a957
15 McHugh GA, Campbell M, Luker KA. Quality of care for individuals with osteoarthritis: a longitudinal study. *J Eval Clin Pract* 2012;18:534–41.
16 Campbell SM, Braspenning J, Hutchinson A, *et al*. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002;11:358–64.
17 Lawrence M,, Olesen F. Indicators of Quality in Health Care. *Eur J Gen Pract* 1997;3:103–8.
18 Hochberg MC. Quality measures in osteoarthritis. *Clin Exp Rheumatol* 2007;25(6 Suppl 47):102–6.
19 Strömbeck B, Petersson IF, Vliet Vlieland TPM. Health care quality indicators on the management of rheumatoid arthritis and osteoarthritis: A literature review. *Rheumatology* 2013;52:382–90.
20 Hunter DJ, Neogi T, Hochberg MC. Quality of osteoarthritis management and the need for reform in the US. *Arthritis Care Res* 2011;63:31–8.
21 Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking systematic reviews in health care*. York: University of York, 2009:292.
22 Agency for Healthcare Research and Quality. http://www.ahrq.gov
23 University of Oxford: Centre for Evidence Based Medicine. Levels of Evidence 2009. http://www.cebm.net/index.aspx?o=1025
24 Campbell SM, Cantrill JA. Consensus methods in prescribing research. *J Clin Pharm Ther* 2001;26:5–14.
25 Geraedts MAX, Selbmann H-K, Ollenschlaeger G. Critical appraisal of clinical performance measures in Germany. *Int J Qual Health Care* 2003;15:79.
26 Marshall MN, Shekelle PG, McGlynn EA, *et al*. Can health care quality indicators be transferred between countries? *Qual Saf Health Care* 2003;12:8–12.
27 Wollersheim H, Hermens R, Hulscher M, *et al*. Clinical indicators: development and applications. *Neth J Med* 2007;65:15–22.
28 Boers M, Brooks P, Strand CV, *et al*. The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998;25:198–9.
29 Kerr E, Asch S, Hamilton E, *et al*. *Quality of care for general medical conditions: a review of the literature and quality indicators*. Santa Monica, CA: RAND Health, 2000.
30 McGlynn E, Asch S, Adams J, *et al*. The quality of health care delivered to adults in the United States. *New Engl J Med* 2003;348:2635.
31 Asch S, McGlynn E, Hogan M, *et al*. Comparison of Quality of Care for Patients in the Veterans Health Administration and Patients in a National Sample. *Ann Intern Med* 2004;141:938.
32 MacLean C. Quality indicators for the management of osteoarthritis in vulnerable elders. *Ann Intern Med* 2001;135:711.
33 Shekelle P, MacLean C, Morton S, *et al*. Assessing Care of Vulnerable Elders: Methods for Developing Quality Indicators. *Ann Intern Med* 2001;135(8 Part 2):647.
34 Wenger N, Solomon D, Roth C, *et al*. The Quality of Medical Care Provided to Vulnerable Community-Dwelling Older Patients. *Ann Intern Med* 2003;139:740.
35 Chodosh J, Solomon D, Roth C, *et al*. The quality of medical care provided to vulnerable older patients with chronic pain. *J Am Geriatr Soc* 2004;52:756.
36 Higashi T, Shekelle P, Solomon D, *et al*. The Quality of Pharmacologic Care for Vulnerable Older Patients. *Ann Intern Med* 2004;140:714.
37 Higashi T, Shekelle P, Adams J, *et al*. Quality of Care Is Associated with Survival in Vulnerable Older Patients. *Ann Intern Med* 2005;143:274.
38 Ganz D, Chang J, Roth C, *et al*. Quality of osteoarthritis care for community-dwelling older adults. *Arthritis Rheum* 2006;55:241.
39 MacLean C, Louie R, Shekelle P, *et al*. Comparison of administrative data and medical records to measure the quality of medical care provided to vulnerable older patients. *Med Care* 2006;44:141.
40 Østerås N, Garratt A, Grotle M, *et al*. Patient-reported quality of care for osteoarthritis: development and testing of the osteoarthritis quality indicator questionnaire. *Arthritis Care Res* 2013;65:1043–51.
41 Steel N, Melzer D, Shekelle P, *et al*. Developing quality indicators for older adults: transfer from the USA to the UK is feasible. *Qual Saf Health Care* 2004;13:260.
42 Saliba D, Solomon D, Rubenstein L, *et al*. Quality indicators for the management of medical conditions in nursing home residents. *J Am Med Dir Assoc* 2005;6:S36.
43 Cadogan M, Schnelle J, Al-Sammarrai N, *et al*. A standardized quality assessment system to evaluate pain detection and management in the nursing home. *J Am Med Dir Assoc* 2005;6:1.
44 Zingmond DS, Saliba D, Wilber KH, *et al*. Measuring the quality of care provided to dually enrolled medicare and medicaid beneficiaries living in nursing homes. *Med Care* 2009;47:536.
45 Smith K, Soriano T, Boal J. Brief communication: National quality-of-care standards in home-based primary care. *Ann Intern Med* 2007;146:188.
46 Wenger N, Roth C, Shekelle P, *et al*. Introduction to the Assessing Care of Vulnerable Elders-3 quality indicator measurement set. *J Am Geriatr Soc* 2007;55:S247.
47 ACOVE Investigators. Assessing Care of Vulnerable Elders—3: Quality Indicators. *J Am Geriatr Soc* 2007;55(Suppl 2):S464.
48 MacLean C, Pencharz J, Saag K. Quality indicators for the care of osteoarthritis in vulnerable elders. *J Am Geriatr Soc* 2007;55(Suppl 2):S383.
49 Marshall M, Campbell S, Hacker J, *et al*. eds. *Quality indicators for general practice*. London: RSM Press, 2002.
50 Kirk S, Campbell S, Kennell-Webb S, *et al*. Assessing the quality of care of multiple conditions in general practice: practical and methodological problems. *Qual Saf Health Care* 2003;12:421.
51 MacLean C, Saag KS, Solomon D, *et al*. Measuring quality in arthritis care: methods for developing the Arthritis Foundation's quality indicator set. *Arthritis Rheum* 2004;51:193.
52 Pencharz J, MacLean C. Measuring quality in arthritis care: the Arthritis Foundation's Quality Indicator set for osteoarthritis. *Arthritis Rheum* 2004;51:538.
53 Li L, Sayre E, Kopec J, *et al*. Quality of nonpharmacological care in the community for people with knee and hip osteoarthritis. *J Rheumatol* 2011;38:2230.
54 American Academy of Orthopaedic Surgeons, Physician Consortium for Performance Improvement. Osteoarthritis: Physician Performance Measurement Set. 2006. http://www.aaos.org/research/committee/evidence/PCPI_Measures_Osteoarthritis.pdf
55 EUMUSC.net. Health care quality indicators for OA. 2012. http://www.eumusc.net/workpackages_wp6.cfm
56 Higgins JPT, Green S. eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011.
57 House of Commons Committee of Public Accounts. *Department of Health: Prescribing costs in primary care*. London: The Stationery Office Limited, 2007.
58 NHS Employers. Quality & Outcomes Framework. 2013. http://www.nhsemployers.org/PayAndContracts/GeneralMedicalServicesContract/QOF/Pages/QualityOutcomesFramework.aspx

Appendix B. Peer-reviewed publication resulting from the template development and implementation

RHEUMATOLOGY

Rheumatology 2015;54:844–853
doi:10.1093/rheumatology/keu411
Advance Access publication 20 October 2014

Original article

Quality of care for OA: the effect of a point-of-care consultation recording template

John J. Edwards[1], Kelvin P. Jordan[1], George Peat[1], John Bedson[1], Peter R. Croft[1], Elaine M. Hay[1] and Krysia S. Dziedzic[1]

Abstract

Objective. The aims of this study were to determine the feasibility of introducing a computerized template for identifying quality of care during an OA consultation, describe quality of OA care in practices in which the template was introduced and assess the effect of the template on routinely recorded clinician behaviour in those practices.

Methods. A computerized template to assist the recording of care in consultations for patients with OA was installed in eight general practices. Eligible patients were those $\geqslant$45 years of age consulting for clinical OA during a 6 month period. The main outcomes were frequency of template triggering, achievement of quality indicators during the consultation (assessment of pain and function, assessment for first-line analgesics, provision of information, exercise advice, consideration of physiotherapy referral, weight loss advice) and change in routinely recorded clinician behaviour (diagnostic coding, prescribing, referral, use of radiography, weight records) compared with the 12 months prior to template installation.

Results. The template was triggered for 1730 patients. Achievement of indicators ranged from 36% (for consideration of physiotherapy referral) to 63% (for pain assessment), with substantial variability between clinicians. There was an increase in prescription of recommended first-line analgesics following the template installation: paracetamol [odds ratio (OR) 1.49 (95% CI 1.22, 1.82) compared with pre-template] and topical NSAIDs [OR 1.95 (95% CI 1.61, 2.35)].

Conclusion. This new template is a feasible tool for capturing data during OA consultations to aid assessment of quality of care. It was associated with significant improvements in recommended care processes. However, strategies are needed to ensure consistent approaches between clinicians.

Trial registration. http://www.controlled-trials.com/ISRCTN06984617/mosaics.

Key words: osteoarthritis, primary care, quality indicator, reminder systems, medical record systems, computerized.

CLINICAL SCIENCE

Introduction

OA is a leading cause of disability: the Global Burden of Disease 2010 ranked OA 11th in the global causes of years lived with disability [1]. A recent review of the UK's health performance concluded that 'interventions are available for musculoskeletal disorders, but to what extent the health system is delivering is unclear' [2]. Guidelines recommend a range of evidence-based treatment options for OA [3–8], and yet European and other surveys have demonstrated suboptimal management compared with guideline recommendations, including underuse of non-pharmacological measures, including exercise and weight loss, and suboptimal pharmacological management [9–13]. Most health care contacts for OA occur within primary care. In the UK, 4% of adults aged $\geqslant$45 years consult for diagnosed OA each year, with the prevalence rising with age [14]. This equates to more than a million people in the UK consulting primary care for OA in a year, and 8.75 million people in the UK have sought treatment for OA [15]. Although there are no

[1]Arthritis Research UK Primary Care Centre, Research Institute for Primary Care and Health Sciences, Keele University, UK

Submitted 22 January 2014; revised version accepted 21 August 2014

Correspondence to: John J. Edwards, Arthritis Research UK Primary Care Centre, Research Institute for Primary Care Sciences, Keele University, Keele ST5 5BG, UK. E-mail: j.j.edwards@keele.ac.uk

agreed benchmarks for performance, there is a recognized need to improve many aspects of primary care for OA [16].

Quality of care in general practice is generally assessed using process of care measures [17]. For OA, these could relate to patient assessment, investigation, information provision, pharmacological and non-pharmacological management and referral [18]. The indicators most feasibly implemented depend on prescribing data, which in the UK is generally electronically recorded and easily audited. There are difficulties with routine use of other potential indicators due to problems with identification of those receiving the care process (numerator) and those eligible for such care (denominator). For example, the need for and use of investigations and referral are not consistently well-captured by the primary care electronic record.

Computerized templates or point-of-care reminders have been shown to have small to moderate effects in improving the quality of consultations [19–22]. This may be due partly to better recording, but it has also been attributed to improved processes of care [19]. A trend has been identified toward greater effects for reminders that require an active response from the clinician [22]. Oliver [23] described a template for the multidisciplinary assessment of OA and RA, though there is a lack of evidence to describe the implementation and effect of computer templates in the management of OA.

The objectives of this study were, through a novel implementation of some of the principles of computerized templates, to determine the feasibility of introducing such a template for identification of quality of care during an OA consultation, describe quality of care for OA consultations in practices in which the template was introduced and assess the effect of the template on clinician behaviour, including pharmacological and some non-pharmacological aspects of management.

Methods

This study was in two parts. The first was an assessment of quality of care for OA in primary care using data collected through a new point-of-care consultation recording template over a 6 month period. The second was a before-and-after study using routinely recorded management actions as a means to estimate the effect of the template on the management of OA in primary care. The study was nested within a wider research programme [the Management of Osteoarthritis in Consultations (MOSAICS) study] designed to investigate effective ways to implement national guidelines for primary care treatment of OA [3].

The study was set in eight general practices with validated data quality in the West Midlands and North West of England that varied in the size of the patient population, clinical staffing, urbanization and local deprivation [24]. The practices received funding for additional costs of participation, dependent upon their expected consultation prevalence for OA but not upon study performance.

A computerized template to record management during an OA consultation for use in general practices was developed (supplementary Fig. S1, available at *Rheumatology* Online). The content of the template was determined from a systematic review of quality indicators for the primary care of OA [18]. The quality indicators related to aspects of OA management unlikely to be captured in medical records and reflected aspects of the UK National Institute of Health and Care Excellence (NICE) 2008 guidelines for the management of OA [3]. The indicators are shown in Table 1, together with predetermined response options and criteria for achievement. The template also facilitated the entry of weight measurements to calculate BMI. Clinicians could enter data contemporaneously throughout a consultation or complete the template at the consultation end. The clinician could record entries for all the template, for selected parts or bypass the entire template.

The template was triggered by entry of an OA code, or selected joint pain codes considered to represent a working diagnosis of OA, for patients $\geqslant$45 years of age consulting at the practice, by telephone or by home visit in the 6 months after template installation. In UK primary care, morbidities are entered using the Read system of coding. Our previous work demonstrated that clinicians use Read codes in $>$95% of all consultations [24]. Relevant joint pain codes from the Read hierarchy were determined by a panel of six general practitioners (GPs) with an interest in musculoskeletal conditions. The template was tested for practicality in two non-study practices prior to the study. Training was provided to all clinicians in the participating practices at the time of template installation (June–August 2011). This consisted of a meeting between an academic GP from the study team and the GPs and practice nurses in the study practices. Although the wider MOSAICS study context was explained to practices, this was a brief general overview only and there was no inclusion of OA management advice or training. In orientating clinicians to the template, there was an emphasis on routine OA management and on restricting use of the template to improve recording of aspects of that routine clinical practice that were considered relevant by the clinicians. It was made clear that clinicians could fill in only those aspects considered appropriate and that the whole template could be bypassed if not considered relevant for a particular patient. A paper copy of the slide presentation and supplementary explanatory DVD were provided for future reference and to facilitate a cascade of training to other team members if required. These supplementary materials were confined to explanation of the use of the template as a recording tool. Neither practices nor clinicians were provided with copies of the NICE OA management guidance, nor were these presented or otherwise reinforced. After 3 months of use, an interim analysis of template data was undertaken to ensure that the template was triggering as expected and that associated data were captured. Feedback sessions between the practices and investigators were held after the interim analysis, but no changes were made to the template. The frequency of template

TABLE 1 Quality indicators included in the template, response options and criteria for achievement

Quality indicator	Response options	Criterion achieved if recorded as	Criterion not achieved if
Pain assessment	None Mild Moderate Severe	None or mild or moderate or severe	No entry
Functional limitation assessment	None Mild Moderate Severe	None or mild or moderate or severe	No entry
Topical NSAID use	Tried full dose Offered full dose Patient declined full dose Not appropriate Unknown	Tried full dose or offered full dose or patient declined full dose or not appropriate	No entry or unknown
Paracetamol use	Tried full dose Offered full dose Patient declined full dose Not appropriate Unknown	Tried full dose or offered full dose or patient declined full dose or not appropriate	No entry or unknown
OA information given	Verbal and written Verbal only Not appropriate Not this time	Verbal and written or verbal only or not appropriate	No entry or not this time
Weight loss advice[a]	Verbal and written Verbal only Not appropriate Not this time	Verbal and written or verbal only or not appropriate	No entry or not this time
Exercise advice	Verbal and written Verbal only Not necessary Not appropriate Not this time	Verbal and written or verbal only or not necessary or not appropriate	No entry or not this time
Consideration of physiotherapy referral	Offered Not necessary Not appropriate Not this time	Offered or not necessary or not appropriate	No entry or not this time

[a]In those with a recorded BMI $\geqslant 25$ in the previous 3 years.

triggering was used as an indicator of the feasibility of template use.

Data entered through the template during the 6 months after installation were used to assess achievement of quality indicators for the care of OA in all patients whose consultation triggered the OA template. We ascertained whether each indicator on the template had been achieved for a patient at any time during the 6 months. The weight advice indicator was only assessed in overweight patients (with a most recent BMI record in the previous 3 years of $\geqslant 25\,kg/m^2$). We identified the first (index) clinician to enter a relevant OA or joint pain code for each patient during the 6 month observation period.

Changes in clinician behaviour were assessed separately from the template-collected information. We used analysis of management actions, which are routinely recorded outside of the template (see below), enabling a before-and-after template installation comparison of management. Routinely recorded medical records data relating to management actions for OA were extracted for all eligible patients with an OA diagnosis code or selected joint pain code recorded in a consultation during three time periods: (i) 12 to 6 months prior to template installation (period 1), (ii) the 6 months prior to template installation (period 2) and (iii) the 6 months after template installation (period 3). This allowed clinician behaviour in period 3 (post-installation) to be compared with a period of equivalent length immediately pre-installation (period 2), and with the identical calendar period in the previous year (period 1).

Management actions for OA included weight records, prescription data, use of radiographs and referrals, all identified from the electronic medical records within 14 days of an OA or joint pain consultation. Prescriptions for paracetamol, topical NSAIDs, opioids and oral NSAIDs were identified. Prescription data were

independent of the template but the template contained a prompt to clinicians regarding paracetamol and topical NSAIDs. In those who were prescribed oral NSAIDs, we determined whether the patient had been prescribed a proton pump inhibitor. We identified records of weight or BMI, relevant referrals (rheumatology, orthopaedics, pain clinic, physiotherapy, occupational therapist, exercise or weight loss programme) and relevant X-rays (knee, hip, hand or foot). As the template could also prompt clinicians to consider physiotherapy referral, we assessed this separately and jointly with other referrals. If a patient consulted more than once for OA or joint pain during a period, they were counted in the denominator only once but were recorded as having received a management action (the numerator) if they had received it within 14 days of any eligible consultation.

Ethical approval has been granted for this study [North West 1 Research Ethics Committee (Cheshire), reference no. 10/H1017/76].

Statistical analysis

Stability in consultation prevalence of recorded OA and joint pain before and after template installation was assessed to ensure the template did not alter morbidity recording habits. The feasibility of using the template was assessed by whether it successfully fired on entry of a relevant code, how often an entry was made after it had fired and the extent of variability in completion between clinicians.

For each template indicator the percentage of patients with recorded achievement during the 6 month period after installation was determined along with its 95% CI, accounting for clustering by practice. We determined the percentage of patients with at least one indicator achieved and with all indicators achieved. For those with a record of being overweight, there were a maximum of eight indicators, otherwise there were seven (excluding weight loss advice). Achievement of indicators was stratified in two ways: (i) by whether the patient was consulting for a new episode (defined as no recorded consultation for OA or joint pain in the 12 months prior to template installation) and (ii) by whether the patient had been given an OA or a joint pain label. Associations between receiving an OA rather than a joint pain label and indicator achievement were assessed through multilevel logistic regression, accounting for clustering within clinician and adjusting for practice. Similar analysis assessed associations of a new episode with indicator achievement. The analysis was repeated for those patients with at least one recorded entry in the template, on the premise that any template entry implies that patients were more likely to be considered by the clinician as having OA.

The monthly percentage of consultations for OA and joint pain that had each management action recorded was plotted to assess trends over the 18 months. Then the percentage of patients with the recorded management action was compared between the three 6 month time periods. Multilevel logistic regression was used to take into account clustering of patients within clinician. Results are presented as ORs with 95% CIs, using period 1 as the reference category and adjusted for patient age, gender, multiple OA consultations in the same period, whether the patient received an OA or joint pain label and practice. All multilevel models were estimated using iterative generalized least squares with second-order penalized quasi-likelihood approximation. STATA version 12.1 (StataCorp, College Station, TX, USA), MLwiN version 2.26 (Centre for Multilevel Modelling Graduate School of Education, University of Bristol, Bristol, UK) and the STATA command runmlwin were used for the analyses [25, 26].

Results

In the 6 months after installation, the template fired for 1730 (93%) of the 1851 patients with a recorded OA or joint pain code. The template fired once for 1255 patients (73%) and twice for 325 patients (19%), up to a maximum of 10 times. A total of 86 clinical staff fired the template with a median of 14 patients each (range 1–82). The consultation prevalence rate for OA or joint pain for adults aged $\geqslant$45 years in the first 6 months after template installation was 549/10 000 (95% CI 525, 574) [27], similar to estimates derived from consultation data of 12 general practices contributing to our local Consultations in Primary Care Archive consultation database [24, 28] (projected rate 500/10 000) (supplementary Table S1, available at *Rheumatology* Online).

Of the 1730 patients, 1147 (66%) patients had at least one entry on the template, with 1146 [OR 66% (95% CI 54, 79)] having at least one indicator achieved and 352 (20%) having all indicators achieved (Table 2). However, this varied greatly by index clinician: for those triggering the template in $>$14 (median) patients, 26% achieved at least one indicator for $>$88% of their patients. However, another quarter failed to achieve any indicator for more than half of their patients. Pain (63%) and function (62%) assessment indicators were achieved most frequently and consideration of physiotherapy referral the least (36%). The only difference in achievement of individual indicators between new episode and ongoing consulters was for consideration of physiotherapy referral, where a higher percentage of ongoing consulters had evidence of achievement (40% *vs* 34%, $P = 0.001$). However, patients with an OA rather than a joint pain label had higher levels of recorded achievement across the indicators (all $P < 0.05$). Those with an OA label were also more likely to achieve all indicators (28% *vs* 17%, $P < 0.001$; Table 2).

When restricted to the 1147 patients with at least one template entry, indicator achievement ranged from 96% for pain assessment to 54% for consideration of physiotherapy referral (Table 3); OR 31% (95% CI 15, 46) of patients had achievement of all indicators. However, wide variation between clinicians remained. There were differences in achievement between those with an OA label and joint pain for four indicators and for achievement of all indicators (39% *vs* 26%, $P < 0.001$).

The 6 month consultation prevalence of OA and joint pain across the eight practices increased from

Table 2 Achievement of indicators in all patients for whom the template fired ($n = 1730$)

Indicator	Patients achieving indicator[a], *n*	Patients achieving indicator, % (95% CI)	Patients achieving indicator across clinicians[b], IQR (range), %	Patients with OA label achieving indicator ($n = 588$), *n* (%)	Patients with joint pain label achieving indicator ($n = 1142$), *n* (%)	Achievement of indicator (OA label *vs* joint pain label), OR[c] (95% CI)
Pain assessment	1097	63 (50, 77)	42–87 (0–100)	403 (69)	694 (61)	1.34 (1.02, 1.74)
Function assessment	1075	62 (50, 75)	50–85 (0–100)	392 (67)	683 (60)	1.30 (1.00, 1.69)
Topical NSAID use	835	48 (35,61)	25–73 (0–100)	313 (53)	522 (46)	1.45 (1.13, 1.88)
Paracetamol use	974	56 (43, 69)	38–79 (2–98)	368 (63)	606 (53)	1.50 (1.17, 1.92)
Information given	852	49 (33, 66)	33–76 (0–100)	339 (58)	513 (45)	1.62 (1.26, 2.09)
Exercise advice	867	50 (35, 65)	27–76 (2–100)	345 (59)	522 (46)	1.60 (1.25, 2.05)
Physiotherapy referral	618	36 (22, 50)	19–53 (0–88)	244 (41)	374 (33)	1.32 (1.03, 1.69)
Weight loss advice[d]	484	45 (27, 62)	20–64 (5–94)	200 (52)	284 (41)	1.38 (1.01, 1.90)
At least one indicator achieved	1146	66 (54, 79)	50–88 (5–100)	416 (71)	730 (64)	1.32 (1.01, 1.73)
All indicators achieved	352	20 (7, 33)	6–32 (0–78)	163 (28)	189 (17)	1.87 (1.38, 2.52)

[a]Number of patients with a record of indicator achievement; [b]in those clinicians with at least 14 patients; [c]adjusting for practice and accounting for clustering by clinician; [d]in those recorded as overweight: total, $n = 1085$; OA, $n = 386$; joint pain, $n = 699$. IQR: interquartile range; OR: odds ratio.

Table 3 Achievement of indicators in patients with at least one entry on the template ($n = 1147$)

Indicator	Patients achieving indicator[a], *n*	Patients achieving indicator, % (95% CI)	Patients achieving indicator across clinicians[b], IQR (range), %	Patients with OA achieving indicator ($n = 416$), *n* (%)	Patients with joint pain achieving indicator ($n = 731$), *n* (%)	Achievement of indicator (OA label *vs* joint pain label)[c], OR (95% CI)
Pain assessment	1097	96 (93, 99)	94–100 (0–100)	403 (97)	694 (95)	—
Function assessment	1075	94 (90, 98)	86–100 (0–100)	392 (94)	683 (93)	—
Topical NSAID use	835	73 (60, 86)	50–96 (0–100)	313 (75)	522 (71)	1.55 (1.05, 2.28)
Paracetamol use	974	85 (78, 92)	78–100 (43–100)	368 (88)	606 (83)	1.59 (1.09, 2.32)
Information given	852	74 (59, 89)	50–94 (0–100)	339 (81)	513 (70)	1.91 (1.32, 2.76)
Exercise advice	867	76 (63, 88)	57–91 (19–100)	345 (83)	522 (71)	1.80 (1.25, 2.57)
Physiotherapy referral	618	54 (40, 68)	33–64 (0–100)	244 (59)	374 (51)	1.30 (0.97, 1.73)
Weight loss advice[d]	484	64 (48, 80)	36–87 (8–100)	200 (70)	284 (60)	1.32 (0.89, 1.96)
All indicators achieved	352	31 (15, 46)	12–42 (0–100)	163 (39)	189 (26)	1.95 (1.41, 2.71)

[a]Number of patients with a record of indicator achievement; [b]in those clinicians with at least 14 patients; [c]adjusting for practice and accounting for clustering by clinician; [d]in those recorded as overweight: total, $n = 758$; OA, $n = 287$; joint pain, $n = 471$. IQR: interquartile range; OR: odds ratio.

TABLE 4 Number of people consulting for OA or joint pain per 10 000 people aged ⩾45 years in each 6 month period

Time period	OA prevalence per 10 000 (95% CI)	Joint pain prevalence per 10 000 (95% CI)	OA or joint pain prevalence per 10 000 (95% CI)
6–12 months before template introduced	174 (161, 189)	384 (363, 405)	522 (498, 547)
0–6 months before template introduced	192 (177, 207)	387 (366, 408)	542 (518, 568)
0–6 months after template introduced	182 (168, 197)	392 (371, 414)	549 (524, 574)
Change[a], %	5	2	5

[a]Percentage change in consultation prevalence from 6–12 months before template to 0–6 months after template introduction.

522/10 000 registered population to 549/10 000 from periods 1 to 3, but the majority of this increase occurred between periods 1 and 2 (before template installation) (Table 4). One practice increased prevalence by 33% between periods 1 and 3 and another by 18%. Four practices increased prevalence by ⩽2% (Appendix 2). Comparison between the three periods showed no change in the likelihood of recording an OA rather than a joint pain label after template installation [period 3 *vs* 1: OR 1.01 (95% CI 0.86, 1.18)], or in age or gender distribution of patients.

A total of 4412 people consulted for OA or joint pain during the three periods of the study and 3511 (80%) of these only consulted in one of the three periods. In the various periods, 90–94 clinicians at the eight practices saw a median of 12–14 patients (range 1–82) with OA or joint pain.

Fig. 1 shows monthly trends in routinely recorded management actions. The percentage of patients receiving each action did not change between periods 1 and 2. Between periods 2 and 3, i.e. before and after template installation, there was a significant change only in management actions, which were also subject to recording prompts in the template: OR 1.95 (95% CI 1.61, 2.35) for topical NSAID prescription; OR 1.49 (95% CI 1.22, 1.82) for paracetamol prescription; OR 3.38 (95% CI 2.73, 4.19) for a weight record (Table 5). The increase in topical NSAID and paracetamol prescribing led to a smaller increase in prescribing of any analgesic between periods 1 and 3 [OR 1.35 (95% CI 1.17, 1.57)]. However, there was no increase in physiotherapy referral rates, which was also prompted for consideration on the template.

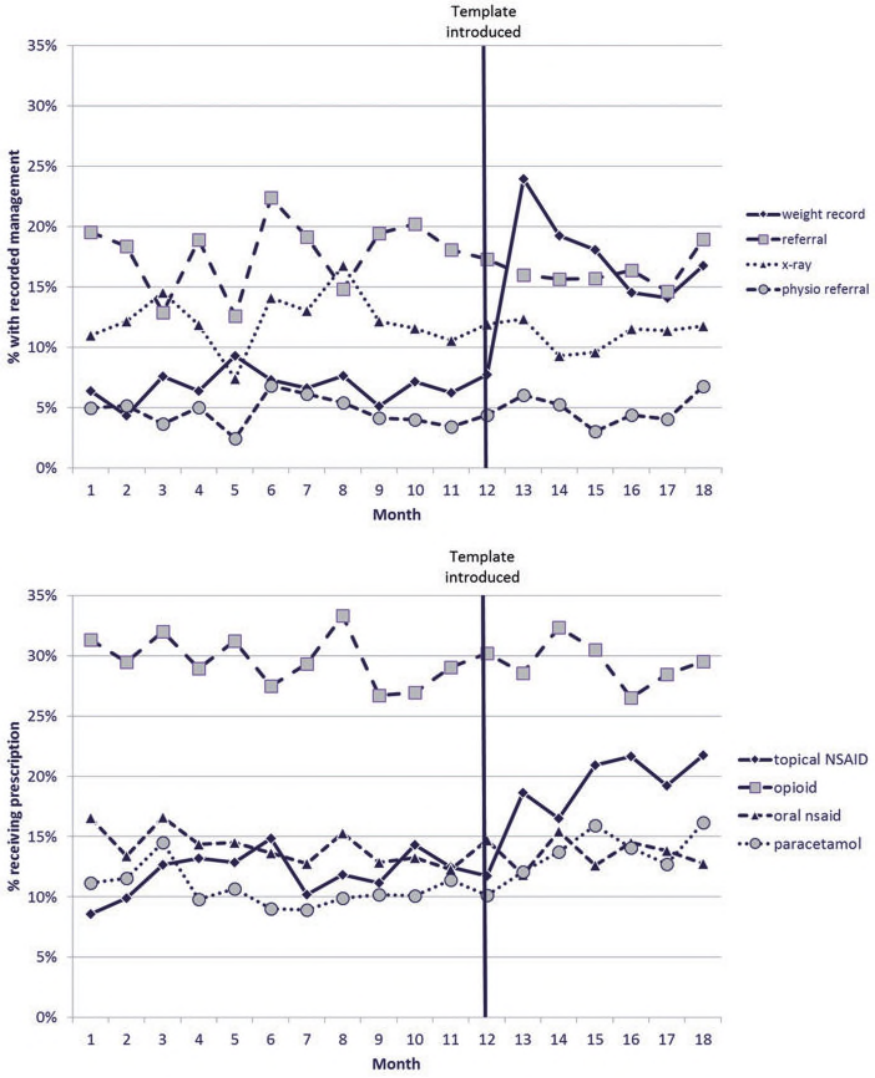
Discussion

Our study found that the principles of a computerized recording template for OA could feasibly be implemented. The general practice staff accepted the template as part of their routine work and the template triggered on 93% of expected occasions. Morbidity coding of OA and joint pain remained stable after the template was introduced, suggesting that clinicians were not avoiding the template through a change in coding behaviour. This was confirmed by the observation that the proportion of people recorded with an OA or joint pain code in the 6 months after template introduction was rather higher than expected. Although there was variation in the way clinicians completed the template, the best-performing clinicians achieved high rates of template completion and quality indicator achievement. The inclusion of prompts to consider the recommended first-line analgesics (topical NSAIDs and paracetamol) also led to an increase in their prescribing.

Despite variability between clinicians, this study has demonstrated greater levels of quality achievement using only electronically coded information than previous studies in UK general practice, which used both electronic and paper records. For example, achievement of pain (63%) and function (62%) assessment indicators compares favourably with rates of 27% and 43% in Broadbent *et al*. [29]. Assessment of indicators for first-line analgesics showed higher rates of achievement than previously reported: advice about first use of paracetamol was 56% in our study compared with 41% in Steel *et al*. and 48% by Broadbent *et al*. [29, 30].

Higher levels of quality achievement were shown when at least one item in the recording template was completed. It is feasible that some patients given a joint pain code were not considered by the recording clinician to have OA and hence the entire template was skipped. More than a quarter of patients given an OA label did not have any template entry. It is plausible that some of these would have achieved some indicator of quality of care, so the actual quality of care delivered may be slightly higher than the recorded level shown. The difference between the achievement rates of patients with an OA label and one of joint pain is partly explained by the joint pain codes' lack of specificity for OA. However, even in patients in whom a template entry had been made (and thus might be considered to have a working diagnosis of OA), the overall recorded quality of care for diagnosed OA was better than that for patients with a joint pain label. There may be a perceived difference in disease in patients with an OA label, or it may be that those clinicians more likely to make a formal diagnosis of OA are also more likely to adhere to guidelines. There may be an order effect, as pain and function assessment were the two most commonly completed entries as they were at the start of the template. Other indicators were less frequently

FIG. 1 Management occurring within 14 days of consultation for OA or joint pain by month.



achieved, with consideration of physiotherapy referral the least frequent.

The template, which reminded clinicians to consider recommended first-line pharmacological and non-pharmacological treatments, resulted in a modest increase in prescriptions of paracetamol and topical NSAIDs but not in physiotherapy referral rates. There was no effect on other interventions: prescription of opioids, oral NSAIDs, proton pump inhibitors, referrals in general or X-ray requests.

Prescriptions for paracetamol increased from 13% to 17% of patients and topical NSAIDs increased from 15% to 25%. The proportion of patients prescribed any analgesic increased after template installation. This increase is greater than might be expected from temporal trends alone [31]. None of the management actions increased in the 12 months prior to template installation, suggesting these changes were not due to temporal factors. Since questions relating to assessment or advice about paracetamol and topical NSAIDs are contained within the template, the template appears to have acted as a prompt for pharmacological management of OA. The heterogeneous nature of reminders, templates and decision support tools as interventions makes direct comparison with other studies unreliable, although these prescribing changes would be consistent with the effects reported in two systematic reviews of computer reminders [19, 22]. The management of several long-term conditions has been found to be improved through the use of reminders and templates, including assessment of cardiovascular disease risk [32]. Computer-guided consultations have also been found to improve aspects of chronic obstructive pulmonary disease management in primary care [33]. In diabetes care, computerized decision support was

TABLE 5 Characteristics of consulters for OA or joint pain and management actions within 14 days of consultation by period

	Period 1: 6–12 months before template	Period 2: 0–6 months before template	Period 3: 0–6 months after template	Period 2 *vs* period 1, OR[a] (95% CI)	Period 3 *vs* period 1, OR[a] (95% CI)
Consulters[b], *n*	1760	1829	1851		
Female, *n* (%)	1035 (59)	1131 (62)	1090 (59)		
Age, mean (S.D.), years	66.2 (11.79)	66.4 (11.79)	66.1 (11.90)		
OA diagnosis[c], *n* (%)	588 (33)	646 (35)	614 (33)	1.05 (0.90, 1.23)	1.01 (0.86, 1.18)
Prescriptions, *n* (%)					
Paracetamol	234 (13)	231 (13)	319 (17)	0.94 (0.76, 1.15)	1.49 (1.22, 1.82)*
Topical NSAID	270 (15)	275 (15)	461 (25)	0.96 (0.79, 1.17)	1.95 (1.61, 2.35)*
Opioids	573 (33)	600 (33)	588 (32)	1.02 (0.88, 1.18)	1.00 (0.86, 1.17)
Any analgesic prescription	977 (56)	1032 (56)	1129 (61)	1.04 (0.90, 1.20)	1.35 (1.17, 1.57)*
Oral NSAID	309 (18)	297 (16)	300 (16)	0.90 (0.75, 1.09)	0.90 (0.74, 1.09)
PPI[d]	102 (33)	108 (36)	101 (34)	1.23 (0.86, 1.77)	1.08 (0.75, 1.57)
Other management, *n* (%)					
Weight record	156 (9)	168 (9)	432 (23)	1.04 (0.82, 1.33)	3.38 (2.73, 4.19)*
Referral	401 (23)	417 (23)	372 (20)	1.01 (0.85, 1.19)	0.89 (0.75, 1,06)
Physiotherapy referral	110 (6)	105 (6)	125 (7)	0.90 (0.67, 1.21)	1.14 (0.86, 1.53)
Radiograph	282 (16)	310 (17)	272 (15)	1.09 (0.89, 1.34)	0.95 (0.77, 1.18)

[a]Adjusted for age, gender, coded OA or joint pain, more than one consultation for OA or joint pain during period and practice and accounting for clustering by clinician. Period 1 is the reference; [b]consultation for OA or joint pain in the period; [c]on the date of consultation; [d]in those prescribed oral NSAIDs during the same 14 day period. *$P < 0.05$. OR: odds ratio; PPI: proton pump inhibitor.

associated with improved processes of care, although patient outcomes only improved when performance feedback or case management was added to computerized decision support [34].

Increasing concern regarding the safety of paracetamol as a first-line analgesic option, coupled with long-standing concerns regarding oral NSAIDs and a wider need to reform OA management [16], means that there is a pressing requirement for strategies to improve primary care uptake of non-pharmacological management for OA. Our computerized template does not record a shift in practice towards greater use of these interventions, though the extent to which this is related to structural or process factors or lack of uptake by patients is not determined.

The diverse nature of the practices participating in the study, in terms of staffing, patient population size, urbanization and deprivation, suggests that the results should have a good level of generalizability to other UK practices. Furthermore, the participating practices, though research-active, were not selected for any particular characteristics beyond the capacity to participate in the study, and received reimbursement only to cover their additional costs associated with participation.

Our study had some limitations. The analyses did not account for repeated visits by patients in more than one period. However, we found that the majority of patients did not consult in more than one period. A sensitivity analysis further accounting for repeated visits led to convergence problems in the multilevel model, but suggested that the conclusions would not change. Prescriptions may not fall within the 14 day period used for analysis and not all X-rays and referrals are electronically recorded by practices. However, there is no reason to suspect this introduced bias in assessment of the template effect, as recording methods were unlikely to have changed during the study. When considering process of care measures, there are concerns about the extent to which improvements in care as recorded in the medical records reflect improvements in recording rather than the actual care delivered. Our study has shown significant increases in the actual prescription of some analgesic prescriptions.

We conclude that a relatively simple point-of-care on-screen recording template for OA can help address recording deficiencies in primary care. With wider uptake, such a template would be a useful basis for auditing core OA care. In addition, the template appears to prompt changes in selected aspects of clinician behaviour. Future research aimed at maximizing the benefit from this should focus on the variation in use between clinicians as well its contribution to improved patient-level outcomes.

Rheumatology key messages

- Computerized templates for OA are feasible in practice and help address recording deficiencies.
- OA recording templates are associated with an increase in the proportion of patients receiving first-line analgesics recommended by existing guidelines.
- There remains a need to improve non-pharmacological care for OA.

Acknowledgements

We thank the patients and staff of the general practices who participated in the MOSAICS trial, as well as Rhian Hughes, Tracy Whitehurst, Ian Thomas, Rebecca Foskett and the study facilitators of the Primary Care Research Network, Central England (North Spoke) for their valuable support. Author contributions: J.J.E. designed the computer template tool and selected the indicators of quality of care, provided training and reviews of the template to the participants and drafted and critically revised the article. He is the guarantor. K.P.J. wrote the statistical analysis plan, cleaned and analysed the data and drafted and critically revised the article. G.M.P. provided advice and epidemiological analysis and critically revised the article. J.B., P.R.C. and E.M.H. all provided analytical advice and critically revised the article. K.S.D. is the chief investigator for the MOSAICS study and provided input to the study design and analysis as well as revising the article. All authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: This work was supported by the National Institute for Health Research (NIHR) through an In-Practice Fellowship to J.J.E., an NIHR Clinical Lectureship to J.B., and an NIHR Programme Grant, number RP-PG-0407-10386. This report is independent research arising from an In-Practice Fellowship supported by the National Institute for Health Research. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health. Although the manuscript was submitted to the funders first, they had no influence over the study design, data collection, analysis or interpretation or the decision to submit the manuscript for publication.

Disclosure statement: K.S.D. has been an invited speaker to the European League Against Rheumatism conference and a member of the NICE OA guideline development group. Keele University has received payments and reimbursements of travel and other expenses related to these activities. K.S.D. has no financial relationships with any organizations that may have an interest in the submitted work in the previous 3 years. J.J.E. and J.B. provide general medical services and benefit financially from the Quality & Outcomes Framework of the General Medical Services contract. All other authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Vos T, Flaxman AD, Naghavi M *et al*. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380:2163–96.

2 Murray CJL, Richards MA, Newton JN *et al*. UK health performance: findings of the Global Burden of Disease Study 2010. Lancet 2013;381:997–1020.

3 National Institute for Health and Clinical Excellence. Osteoarthritis: the Care and Management of Osteoarthritis in Adults. London: National Institute for Health and Clinical Excellence, 2008.

4 Hochberg MC, Altman RD, April KT *et al*. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res 2012;64:465–74.

5 Jordan KM, Arden NK, Doherty M *et al*. EULAR recommendations 2003: an evidence based approach to the management of knee osteoarthritis: report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis 2003;62:1145–55.

6 Zhang W, Doherty M, Arden N *et al*. EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2005;64:669–81.

7 Zhang W, Doherty M, Leeb BF *et al*. EULAR evidence based recommendations for the management of hand osteoarthritis: report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2007;66:377–88.

8 Zhang W, Moskowitz RW, Nuki G *et al*. OARSI recommendations for the management of hip and knee osteoarthritis, part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;16:137–62.

9 Chevalier X, Marre JP, de Butler J *et al*. Questionnaire survey of management and prescription of general practitioners in knee osteoarthritis: a comparison with 2000 EULAR recommendations. Clin Exp Rheumatol 2004;22:205–12.

10 Denoeud L, Mazieres B, Payen-Champenois C *et al*. First line treatment of knee osteoarthritis in outpatients in France: adherence to the EULAR 2000 recommendations and factors influencing adherence. Ann Rheum Dis 2005;64:70–4.

11 Sarzi-Puttini P, Cimmino MA, Scarpa R *et al*. Do physicians treat symptomatic osteoarthritis patients properly? Results of the AMICA experience. Semin Arthritis Rheum 2005;35(Suppl 1):38–42.

12 Conrozier T, Marre JP, Payen-Champenois C *et al*. National survey on the non-pharmacological modalities prescribed by French general practitioners in the treatment of lower limb (knee and hip) osteoarthritis. Adherence to the EULAR recommendations and factors influencing adherence. Clin Exp Rheumatol 2008;26:793–8.

13 Poitras S, Rossignol M, Avouac J *et al*. Management recommendations for knee osteoarthritis: how usable are they? Joint Bone Spine 2010;77:458–65.

14 Jordan KP, Jöud A, Bergknut C *et al*. International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. Ann Rheum Dis 2014;73:212–8.

15 Arthritis Research UK. Osteoarthritis in General Practice: Data and Perspectives. Chesterfield, UK: Arthritis Research UK, 2013.

16 Hunter DJ, Neogi T, Hochberg MC. Quality of osteoarthritis management and the need for reform in the US. Arthritis Care Res 2011;63:31–8.

17 Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. Int J Qual Health Care 2000;12:281–95.

18 Edwards JJ, Khanna M, Jordan KP *et al*. Quality indicators for the primary care of osteoarthritis: a systematic review. Ann Rheum Dis 2013, Nov 27. doi: 10.1136/annrheumdis-2013-203913. [Epub ahead of print].

19 Shojania KG, Jennings A, Mayhew A *et al*. The effects of on-screen, point of care computer reminders on processes and outcomes of care. Cochrane Database Syst Rev 2009;3:CD001096.

20 Linder JA, Rigotti NA, Schneider LI *et al*. An electronic health record-based intervention to improve tobacco treatment in primary care: a cluster-randomized controlled trial. Arch Intern Med 2009;169:781–7.

21 Davis AM, Cannon M, Ables AZ *et al*. Using the electronic medical record to improve asthma severity documentation and treatment among family medicine residents. Fam Med 2010;42:334–7.

22 Cheung A, Weir M, Mayhew A *et al*. Overview of systematic reviews of the effectiveness of reminders in improving healthcare professional behavior. Syst Rev 2012;1:36.

23 Oliver S. Multidisciplinary disease management in rheumatology. J Prof Nurs 2003;19:137.

24 Porcheret M, Hughes R, Evans D *et al*. Data quality of general practice electronic health records: the impact of a program of assessments, feedback, and training. J Am Med Inform Assoc 2004;11:78–86.

25 runmlwin [program]. Stata module for fitting multilevel models in the MLwiN software package. Bristol, UK: Centre for Multilevel Modelling, University of Bristol, 2011.

26 MLwiN [program]. Version 2.1. Bristol, UK: Centre for Multilevel Modelling, University of Bristol, 2009.

27 Abramson JH. WINPEPI updated: computer programs for epidemiologists, and their teaching potential. Epidemiol Perspect Innov 2011;8:1.

28 Jordan K, Clarke AM, Symmons DP *et al*. Measuring disease prevalence: a comparison of musculoskeletal disease using four general practice consultation databases. Br J Gen Pract 2007;57:7–14.

29 Broadbent J, Maisey S, Holland R *et al*. Recorded quality of primary care for osteoarthritis: an observational study. Br J Gen Pract 2008;58:839.

30 Steel N, Bachmann M, Maisey S *et al*. Self reported receipt of care consistent with 32 quality indicators: national population survey of adults aged 50 or more in England. BMJ 2008;337:a957.

31 Bedson J, Belcher J, Martino OI *et al*. The effectiveness of national guidance in changing analgesic prescribing in primary care from 2002 to 2009: an observational database study. Eur J Pain 2013;17:434–43.

32 Holt TA, Thorogood M, Griffiths F. Changing clinical practice through patient specific reminders available at the time of the clinical encounter: systematic review and meta-analysis. J Gen Intern Med 2012;27:974–84.

33 Angus RM, Thompson EB, Davies L *et al*. Feasibility and impact of a computer-guided consultation on guideline-based management of COPD in general practice. Primary Care Respir J 2012;21:425–30.

34 Cleveringa FG, Gorter KJ, van den Donk M *et al*. Computerized decision support systems in primary care for type 2 diabetes patients only improve patients' outcomes when combined with feedback on performance and case management: a systematic review. Diabetes Technol Ther 2013;15:180–92.

Appendix C. MOSAICS ethics committee approval

NHS

National Research Ethics Service

North West 1 Research Ethics Committee – Cheshire
Research Ethics Office
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0161 625 7821

18 November 2010

Professor Krysia Dziedzic
Arthritis Research UK Professor of Musculoskeletal Therapies
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Staffordshire
ST5 5BG

Dear Professor Dziedzic

Full title of study: **Management of Osteoarthritis in Consultations Study: the development of a complex intervention in primary care (MOSAICS)**
REC reference number: **10/H1017/76**

Thank you for your letters of 15th and 16th November 2010. I can confirm the REC has received the documents listed below as evidence of compliance with the approval conditions detailed in our letter dated 26 October 2010. Please note these documents are for information only and have not been reviewed by the committee.

I can confirm that the additional conditions have now been met.

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Covering Letter		16 November 2010
Covering Letter		15 November 2010
Participant Information Sheet: Appendix 4.b.3.2 MCQ Information Leaflet (Questionnaires and Interviews)	2	26 October 2010
Participant Consent Form: Appendix 4.e.9 Manager and administrator Consent Form	2	16 November 2010
Participant Information Sheet: Appendix 4.d.13 Patient information leaflet MOAC 3	2	26 October 2010
Participant Information Sheet: Appendix 4.d.7 Patient information leaflet Group interview MOAC 2	2	26 October 2010
Participant Consent Form: Appendix 4.e.1.2 - Nurse Consent Form Observation MOAC 2	2	16 November 2010
Participant Consent Form: Appendix 4.e.5 HCP Interview Consent Form	2	16 November 2010
Questionnaire: Appendix 2.3 Population Survey Questionnaire	2	15 November 2010

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

10/H1017/76	**Please quote this number on all correspondence**

Yours sincerely

Miss Shehnaz Ishaq
Committee Co-ordinator

Email: Shehnaz.ishaq@northwest.nhs.uk

Copy to: Dr Mark Porcheret
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Keele
ST5 5BG

Ms June Handy
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Keele
ST5 5BG

Rhian Hughes
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Keele
ST5 5BG

Ms Nemonie Marriott
Stoke-on-Trent PCT
London House
4th Floor
Hide Street
Stoke-on-Trent, ST4 1NF

NHS
National Research Ethics Service
NRES Committee North West - Cheshire
Research Ethics Office
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Tel: 0161 625 7815

23 August 2011

Professor Krysia Dziedzic
Arthritis Research UK Professor of Musculoskeletal Therapies
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Staffordshire
ST5 5BG

Dear Professor Dziedzic

Study title:	**Management of Osteoarthritis in Consultations Study: the development of a complex intervention in primary care (MOSAICS)**
REC reference:	**10/H1017/76**
Amendment number:	**Substantial Amendment 1**
Amendment date:	**05 August 2011**

- The amendment proposes to change the timing of the follow up questionnaires and the Medical Records Reviews.
- It is also proposed to add an evaluation process for the GPs and Practice Nurses involved in delivering the new approach. This primarily involves questionnaires but GPs will be invited to undertake a video-recorded simulated consultation.
- There are also some minor changes to the study personnel.
- There are also some minor changes to some of the study documentation.

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
MOSAICS Study Team Changes and Additions		01 July 2011
Summary of Changes - Appendix A		
Investigator CV	Susan Hill	05 August 2011

Notice of Substantial Amendment (non-CTIMPs)	Substantial Amendment 1	05 August 2011
Covering Letter		05 August 2011
Investigator CV	Robert David Smith	
Protocol	2 - Tracked Changes - Appendix E	01 July 2011
A Guide for People Who Have Osteoarthritis - Appendix 3.4	2	01 July 2011
Osteoarthritis Template - Appendix 1.1	2	01 July 2011
Questionnaire: GP Questionnaire 1 - Appendix 4.a.1	1	01 July 2011
Summary/Synopsis	Medical Record Review Flowchart - Version 2 - Appendix 1	01 July 2011
Participant Information Sheet: Training Evaluation: GP Information Sheet - Appendix 4.a	1	01 July 2011
Protocol	2 - Appendix F	01 July 2011
Questionnaire: Nurse Practitioner/Practice Nurse/Training Evaluation Questionnaire 3	1 - Appendix 4.a.7	01 July 2011
Questionnaire: Nurse Practitioner/Practice Nurse/Training Evaluation Questionnaire 2	1 - Appendix 4.a.6	01 July 2011
Questionnaire: Nurse Practitioner/Practice Nurse Baseline Survey/Training Evaluation Questionnaire 1	1 - Appendix 4.a.5	01 July 2011
Questionnaire: GP Questionnaire 3	1 - Appendix 4.a.3	01 July 2011
Questionnaire: GP Questionnaire 2 - Appendix 4.a.2	1	01 July 2011
Participant Information Sheet: PN Information Sheet - Appendix 4.a.4	1	01 July 2011
Reminder Letter - Appendix 4.a.8	1	01 July 2011
Summary/Synopsis	MOSAICS Consultations Questionnaire MCQ Flowchart - Version 2 Appendix 4b	01 July 2011
Letter of invitation to participant	Letter of Invitation - Questionnaire - Version 1 Appendix 4.a.10	01 July 2011
Participant Consent Form: GP Consent to SP Video	1 - Appendix 4.a.9	01 July 2011
Participant Information Sheet: MCQ Information Leaflet (Questionnaires and Interviews) Appendix 4.b.3.2	3	01 July 2011
Participant Information Sheet: MCQ Information Leaflet (Questionnaires) Appendix 4.b.3	2	01 July 2011
MCQ Follow-up Letter (Keele)	2 - Appendix 4.b.9	01 July 2011
Questionnaire: Consultation for Joint Pain 6 month follow up Questionnaire 3	1 - Appendix 4.b.12	01 July 2011
Questionnaire: Consultation for Joint Pain 12 month follow up Questionnaire 4	1 - Appendix 4.b.13	01 July 2011
Questionnaire: Consultation for Joint Pain Minimum data questionnaire	2 - Appendix 4.b.8	01 July 2011
Questionnaire: Consultation for Joint Pain 3 month follow up Questionnaire 2	2 - Appendix 4.b.4.2	01 July 2011
Questionnaire: Consultation for Joint Pain Questionnaire 1	2 - Appendix 4.b.4.1	01 July 2011
Summary/Synopsis	Flowchart 4.d.5 for Patient Expereinces of MOAC 2 - Version 2	01 July 2011
Letter of invitation to participant	Patient Invitation Letter Individual or Group Interview MOAC 2 - Version 2 Appendix 4.d.6	01 July 2011

Participant Consent Form: Interview Consent Form MOAC 1	2 - Appendix 4.d.4	01 July 2011
Interview Reply Slip	2 - Appendix 4.d.8	01 July 2011
Interview Schedules/Topic Guides	Individual or Patient Group Interview MOAC 2 - Version 2 - Appendix 4.d.9	01 July 2011
Participant Consent Form: Manager & Administrator Consent Form - Appendix 4.e.9	2	01 July 2011
Participant Consent Form: HCP Interview Consent Form (All)	2 - Appendix 4.e.5	01 July 2011
Participant Consent Form: Patient Interview Consent Form MOAC 3	2 - Appendix 4.d.16	01 July 2011
Participant Consent Form: Patient Consent Form Individual or Group Interview MOAC 2	2 - Appendix 4.d.10	01 July 2011
Participant Information Sheet: Patient Information Leaflet - Individual or Group Interview MOAC 2	2 - Appendix 4.d.7	01 July 2011

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

10/H1017/76: Please quote this number on all correspondence

Yours sincerely

Mr Jonathan Deans
Chair

E-mail: shehnaz.ishaq@northwest.nhs.uk

Enclosures: *List of names and professions of members who took part in the review*

Copy to:
Dr Mark Porcheret
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Keele
ST5 5BG

Ms June Handy
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Keele
ST5 5BG

Rhian Hughes
Arthritis Research UK Primary Care Centre
Primary Care Sciences
Keele University
Keele
ST5 5BG

Ms Nemonie Marriott
Stoke-on-Trent PCT
London House
4th Floor
Hide Street
Stoke-on-Trent, ST4 1NF

NRES Committee North West - Cheshire

Attendance at Sub-Committee of the REC meeting on 22 August 2011

Name	*Profession*	*Capacity*
Dr Nick Bronnert	GP Member	Expert
Mr Jonathan Deans	Consultant ENT Surgeon	Expert
Mr Peter Ward	Lay member	Lay

Appendix D. Systematic review appendices (Chapter Two)

Appendix D.1. Data extraction tools for the systematic review

Appendix D.1.1. Data extraction sheet for quality indicator development studies

Data extraction sheet for full-text articles: QI development

General Info

Reviewer Choose an item.

Study ID ☐☐☐

Date of extraction ☐☐☐☐☐☐

Date of publication ☐☐☐☐☐☐

Author _____

_____

Title _____

_____

Source _____

_____

Institution (of _____

1st author) _____

Study Characteristics

Population _____

_____

Age (range) _____ Sex: M ☐ F ☐ Both ☐ Ratio:_____

Comments

Funding/Conflict of Interest

Is the source of funding clear? Yes☐ No☐

Are conflicts of interest transparent/minimised? Yes☐ No☐

Quality Indicator development

Method of evidence collection, assessment & assimilation

[This information will be grouped for families of indicators]

Meta-analysis☐
Systematic Review☐
RCTs☐
Epidemiological studies☐
Expert opinion☐
Other (specify)☐
Referenced (specify)☐

What is the consensus method used in indicator development?

[This information will be grouped for families of indicators]

RAND Appropriateness☐
Delphi☐
Nominal group☐
Other (specify)☐
Referenced (specify)☐
Not stated☐

Are numerator & denominator populations clear? - note any limitations, exclusions Yes☐ No☐

Is there an identified mechanism for keeping the indicators updated especially in response to new evidence Yes☐ No☐

If no updating mechanism stated, have the indicators been produced sufficiently recently to be considered current? Yes☐ No☐

Quality Indicator Data

For each QI, extract the following:

Indicator #	Text of indicator	Indicator clear or not, with terms defined where needed?	Structure, Process, or Outcome-based, with comments	Population applied to [if varies from that already identified, e.g. specified exclusions]	Anticipated health benefit (e.g. reduced gastrointestinal bleeding)	Method of measurement	Is a change to routine UK 1^0 care recording of data required for indicator validity?	Performance thresholds (as per audit terms Standard and Target) if stated, with justification	Outcome Data [rate] (numerator/ denominator)	Indicator grouping (e.g. ACOVE/RAND, Arthritis Foundation)
		Yes☐ No☐					Yes☐ No☐			
		Yes☐ No☐					Yes☐ No☐			
		Yes☐ No☐					Yes☐ No☐			
		Yes☐ No☐					Yes☐ No☐			

Appendix D.1.2. Data extraction sheet for quality indicator implementation studies

Data extraction sheet for full-text articles
General Info

Reviewer Choose an item.

Study ID ☐☐☐

Date of extraction ☐☐☐☐☐☐

Date of publication ☐☐☐☐☐☐

Author _____

_____

Title _____

_____

Source _____

_____

Institution (of _____

1st author) _____

Study Characteristics

Population _____

_____

Age (range) _____ Sex: M ☐ F ☐ Both ☐ Ratio:_____

Comments

Funding/Conflict of Interest

Is the source of funding clear? Yes☐ No☐

Are conflicts of interest transparent/minimised? Yes☐ No☐

Quality Indicator Implementation & Testing

Is the method for indicator selection clear? Yes☐ No☐

Is the method for indicator prioritisation clear? Yes☐ No☐

Is the rationale for indicator selection clear? Yes☐ No☐

Has there been any quality assessment of the indicator by the implementation group? Comment on degree of rigour of assessment Yes☐ No☐

Have routine data sources been tested? Yes☐ No☐

- and found to yield the necessary data? Yes☐ No☐
- in a manner suitable for our study? Yes☐ No☐

Is there evidence of reliability (i.e. do repeated data extractions from the same dataset yield the same results?) Yes☐ No☐

Is there evidence of quality improvement where the indicators have been used repeatedly over time on the same population? (External validity) Yes☐ No☐

Is there evidence of the indicators' sensitivity to change in quality of care? Yes☐ No☐

Quality Indicator Data

For each QI, extract the following:

Indicator #	Text of indicator	Indicator clear or not, with terms defined where needed?	Structure, Process, or Outcome-based, with comments	Population applied to [if varies from that already identified, e.g. specified exclusions]	Anticipated health benefit (e.g. reduced gastrointestinal bleeding)	Method of measurement	Is a change to routine UK 1^0 care recording of data required for indicator validity?	Performance thresholds (as per audit terms Standard and Target) if stated, with justification	Outcome Data [rate] (numerator/ denominator)	Indicator grouping (e.g. ACOVE/RAND, Arthritis Foundation)
		Yes☐ No☐					Yes☐ No☐			
		Yes☐ No☐					Yes☐ No☐			
		Yes☐ No☐					Yes☐ No☐			
		Yes☐ No☐					Yes☐ No☐			

Appendix D.2. Prior recorded quality indicator achievement

Indicator	Kirk 2003	McGlynn 2003	Wenger 2003	Asch 2004	Chodosh 2004	Higashi 2004	Cadogan 2005	Higashi 2005	Ganz 2006	Maclean 2006	Steel 2007	Broadbent 2008	Steel 2008	Zingmond 2009	Li 2011	Østerås 2013 (patient self-report)
Assessment																
Pain	-	57.3[a]	40	80	-	-	-	-	60.6		23 - 30	27	-	-	-	36
Function	-	-	-	-	41	-	-	-	-		42 - 44	43	-	-	6.9[b] - 29.2[c]	24
Weight recording	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
X-ray use	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
Advice	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
Education	-	-	33 - 36	-	-	-	-	-	68.7		29 - 31	30	17.7	-	-	-
Education (patient self-report)	-	-	-	-	-	-	-	-	-		-	-	77.8	-	-	19 – 24[d]
Exercise	-	57.3[a]	-	27	-	-	-	-	-		-	-	-	-	25.2	49
Weight loss	-	-	-	-	-	-	-	-	-		-	-	-	-	25	34
NSAID risk advice	-	-	4	-	-	-	-	4	38.6 - 50.0		14 - 20	17	-	-	-	-
Non-pharmacological management																
Physio referral/exercise "prescription"	-	-	0 - 16	-	-	-	46	16	44		-	-	24.8	-	-	43
OT referral	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
Weight loss programme	-	-	-	-	-	-	-	-	-		-	-	-	-	-	5
Pharmacological management																
Paracetamol first analgesic	45[f]	57.3[a]	43	73	-	79	26	43	58.7		45 - 51	48	41.1	-	-	46
Paracetamol maximum dose before alternative	-	-	33	-	-	-	37	33	-		03-Jun	5	-	-	-	-
Topical NSAID	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
Oral NSAID	45[f]	-	-	-	-	-	-		-		58 – 60[e]	59[e]	-	-	-	46
PPI used with oral NSAIDs	45[f]	-	11	-	-	-	-	11	27.4		-	-	-	26.6	-	-
Opioid/stronger analgesic	-	-	-	-	-	-	-	-	-		-	-	-	-	-	37
Capsaicin	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
Glucosamine	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
Rubefacient	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
Referral once nonsurgical treatment failed	45[f]	-	90	-	-	-	-	90	72.7		90 - 95	90	35.8	-	-	47

No indicator-specific data published

[a]composite of 3 indicators; [b] nonambulatory; [c] ambulatory; [d]depending on component of education; [e]for ibuprofen or COX2 as first used oral NSAID; [f]composite of 4 indicators - individual indicator pass rates 27 - 100

Recorded quality achievement in indicator implementation studies identified in the systematic review

Appendix D.3. Studies excluded from the systematic review

Author [Reference number]	Reasons for exclusion
Abraham, et al. [358]	Guideline adherence study
Alghamdi, et al. [359]	Exercise intervention review
Allen, et al. [360]	Intervention study methods paper
American College of Rheumatology [361]	No OA-specific indicators
Arthritis and Musculoskeletal Alliance [362]	Standards of care not quality indicators.
Askari, et al. [363]	Systematic review of ACOVE indicators but no additional development or implementation
Barlow [364]	Education intervention review
Baumann, et al. [365]	Patient expectations study
Bierma-Zeinstra, et al. [366]	Study of GPs management of OA
Bijlsma [367]	Review of OA management, especially analgesia
Blanda [368]	Review of prescribing issues in the elderly
Bliddal and Christensen [369]	Review of management of OA in the overweight
Bloodworth [370]	Review of opiate prescribing issues
Boyd, et al. [371]	Assessment of guideline applicability in multimorbidity
Brand [372]	Review of evidence into practice for OA hip/knee
Brand [373]	Review of self-management in OA
Brand and Cox [374]	Care pathway development for OA
Brooks and Conaghan [375]	Editorial
Brosseau, et al. [376]	Guidelines for exercise/ manual therapy
Cadogan, et al. [377]	Pain assessment study
Chassany, et al. [378]	Education intervention study, no qi
Cieza, et al. [379]	Core dataset recommendations
Conaghan and Brooks [380]	Review of management of OA
Conrozier, et al. [86]	Guideline adherence study
Cornali, et al. [381]	Assessment of OA
Cox, et al. [382]	Intervention review (of COX-2 drugs)
Curtis and Saag [383]	Review of quality improvement – cites indicators but does not additionally develop of implement them.
DeHaan, et al. [84]	Guideline adherence study
Denoeud, et al. [87]	Guideline adherence study
Doherty and Dougados [384]	Review of OA management
Dominick and Baker [385]	Epidemiology of OA and its treatment
Dominick, et al. [386]	Survey of opioid use
Dooley and Martin [387]	Education article
Duran-Barragan and Russell [388]	Editorial
Emejuaiwe, et al. [389]	Epidemiology of joint replacement
Escobar, et al. [390]	Clinical guideline for criteria for knee replacement
Feldt [391]	Editorial
Ganz, et al. [392]	Intervention study (not using OA indicators)

Author [Reference number]	Reasons for exclusion
Glazier, et al. [393]	Guideline adherence study
Gogovor, et al. [394]	Review of clinical management
Gordon, et al. [395]	Letter
Gorevic [396]	Education article
Gossec, et al. [397]	Clinical guideline for criteria for knee replacement
Graudins and Gazarian [398]	Survey (of celecoxib use)
Hall [399]	Survey of information provision
Higashi, et al. [400]	Secondary epidemiological analysis (number of morbidities and quality of care)
Hochberg [105]	Narrative review of OA quality indicators but no additional development or implementation
Hochberg, et al. [38]	Clinical guidelines
Holden, et al. [401]	Guideline adherence study
Hollenack, et al. [402]	Review of pain management in elderly patients
Holman, et al. [403]	Letter
Hunter [404]	Review of guidelines. No development or implementation of quality indicators.
Hunter, et al. [405]	Review of quality of OA care, including indicators; no additional development or implementation.
Jansen, et al. [406]	Guideline adherence study
Jawad [407]	Guideline adherence study
Jones and Piterman [408]	Intervention study (education)
Jordan, et al. [31]	Clinical guidelines
Jordan, et al. [409]	Survey of OA treatments
Kane, et al. [410]	Guideline [Health Technology Assessment (joint replacement)]
Kartal, et al. [411]	Survey of prescribing for OA
Kirkness, et al. [412]	Review of OA pain & comorbidity
Koehn and Esdaile [413]	Review of patient education and self-management.
Koyama, et al. [414]	Survey of patient self-management activities
Krishnan and Suarez-Almazor [415]	Review of evidence-based rheumatology practice – references an indicator set but no additional development or implementation
Labelle, et al. [416]	Intervention study (medical education)
Lai, et al. [417]	Intervention study (hyaluronic acid)
Laine, et al. [418]	Guideline adherence study
Landsberg, et al. [419]	Survey of COX-2 use in secondary care
Lanza, et al. [420]	Clinical guidelines
Lapane, et al. [421]	Epidemiological study (effect of NSAIDs on gastro-protective medication)
Lee and Katz [422]	Education article
LeLorier, et al. [423]	Guideline adherence study (celecoxib)
Li, et al. [424]	Review of treatments. References quality indicators but no additional development or implementation.
LoBuono [425]	News/opinion article. No OA quality indicators in referenced reports.
Maillefert, et al. [426]	Survey of factors influencing joint replacement
March, et al. [427]	Clinical guideline/standards of care
Marcum, et al. [428]	Survey of pain control in OA
Marks and Allegrante [429]	Guideline adherence study (exercise)
Mazieres, et al. [430]	Guideline adherence study
McHugh, et al. [431]	Qualitative study of quality of care. No quality indicators developed or implemented.

Author [Reference number]	Reasons for exclusion
McHugh, et al. [432]	Longitudinal study of quality of care (implementation study). Three indicators of quality of care were reported - the source and development not cited.
McPherson and Reid [433]	Editorial
Merkle and McDonald [434]	Guideline adherence study (pain management)
Min, et al. [435]	Secondary analysis of data included elsewhere [78] to identify predictors of care; no additional development or implementation of quality indicators.
Min, et al. [276]	Secondary analysis of data included elsewhere [78] to identify associations between multimorbidity and quality of care; no additional development or implementation of quality indicators.
Mitchell and Hurley [436]	Survey of patient reported management and preferences
Morewitz [437]	Review of developments in care (book)
Newman and Harrington [438]	Care pathway review (system redesign)
Oliver [439]	Review of a template for multidisciplinary management of OA
O'Reilly and Doherty [440]	Survey to assess link between knee pain & disability and poor health status & psychological distress
Østerås, et al. [441]	Abstract on development of patient self-report quality indicators [excluded as abstract only]
Osterhaus, et al. [442]	Evaluation of new methods of data collection in community pharmacies.
Pendleton, et al. [443]	Clinical guidelines
Peter, et al. [444]	Abstract on development of quality indicators for physical therapy [excluded as abstract only]
Pitt, et al. [445]	Survey of GPs evaluating referral to self-management programmes
Porcheret, et al. [446]	Care pathway development
Price-Forbes, et al. [447]	Audit of COX-2 drugs
Quintana, et al. [448]	Development of criteria for total hip replacement.
Quintana, et al. [449]	Survey of patients undergoing hip replacement to evaluate appropriateness
Quintana, et al. [450]	Survey of patients undergoing hip replacement to evaluate appropriateness
Quintana, et al. [451]	Clinical guideline (joint replacement decision tree)
Reeves, et al. [452]	Secondary analysis (of differences in quality indicator aggregation methods)
Rosemann, et al. [453]	Guideline adherence study
Saag, et al. [454]	Analgesics quality indicators considered too broad for inclusion in this review.
Saliba, et al. [455]	Feasibility assessment of various quality indicators but not those for OA
Sarzi-Puttini, et al. [85]	Guideline adherence study
Schilling [456]	Review of pain management which references another (included) paper but no indicators developed or implemented within this paper.
Shrank, et al. [138]	Analgesics quality indicators considered too broad for inclusion in this review.
Singh, et al. [457]	Review of OA drug treatment
Solomon, et al. [458]	Survey of medical record data, finding that history and examination data as recorded were not suitable for use as quality markers.
Strand [459]	Quality indicators for rheumatoid disease not OA
Strombeck, et al. [106]	Systematic review of quality indicators in OA but no additional development or implementation of indicators
Sturkenboom, et al. [460]	Epidemiological study of gastro-protective strategies in patients receiving NSAIDs.
Vandenberghe, et al. [461]	Comparative study of data sources. Although this does use 5 prescribing indicators, we considered this to be a study investigating the use of limited vs.

Author [Reference number]	Reasons for exclusion
	complete datasets from general practice rather than a development or implementation of the quality indicators *per se.*
Vandenberghe, et al. [462]	Comparative study of data extraction methods
Webster [463]	Editorial
Wierenga, et al. [464]	Pharmaceutical care quality indicators for hospital in-patients (i.e. secondary care not primary care)
Wilcock, et al. [465]	Audit of NSAID prescribing
Yazdany and MacLean [466]	Review of quality of care but no additional development or implementation of indicators
Zhang, et al. [32]	Clinical guidelines
Zhang, et al. [33]	Clinical guidelines
Zhang, et al. [36]	Clinical guidelines
Zhang, et al. [35]	Clinical guidelines
Zhang, et al. [37]	Clinical guidelines
Zingmond, et al. [467]	No OA quality indicators included in this study of quality of care.

Appendix D.4. Holistic approach to OA assessment and management indicators and example synthesis

Development/ Implementation	***Author***	***Date***	***Topic***	***Indicator***	***Extracted theme***		
					Target population	*Process or outcome*	*Timeframe*
Development	Kerr[135]	2000	Pain & function assessment	Providers caring for patients with symptoms of OA should document at least one of the following at least once in 2 years: a. the location of symptoms; b. the presence or absence of limitations in daily activities; c. the use and effectiveness of treatment modalities.	Patients with symptomatic OA…	…should be assessed for symptoms, function, and the effects of treatment…	…every two years.
Development	MacLean[134]	2001	Pain & function assessment	IF a vulnerable elder is diagnosed with symptomatic osteoarthritis, THEN his or her functional status and the degree of pain should be assessed annually BECAUSE this information is necessary to direct therapeutic decisions.	Patients with symptomatic OA…	…should be assessed for pain and function…	…annually.
Development	MacLean[134]	2001	Differential diagnosis assessment	IF a vulnerable elder has monoarticular joint pain associated with redness, warmth, or swelling AND the patient also has an oral temperature greater than 38.0 °C and does not have a previously established diagnosis of pseudogout or gout, THEN a diagnostic aspiration of the painfully swollen red joint should be performed that day BECAUSE this sign–symptom complex is common with joint infection, and it requires treatment that is different than that for osteoarthritis.	Patients with a possible differential diagnosis of a septic joint…	…should have a joint aspiration…	…the same day.
Development	MacLean[127,128]	2004	Examination	IF a patient is begun on a drug treatment for joint pain, arthritis, or arthralgia, THEN evidence that the affected joint was examined should be documented	Patients with joint pain who receive drug treatment…	…should have documentary evidence of a joint examination.	-

Development	MacLean[127,128]	2004	Pain assessment	IF a patient is diagnosed with symptomatic osteoarthritis of the knee or hip, THEN his or her pain should be assessed annually and when new to a practice.	Patients with symptomatic OA (hip or knee)…	…should be assessed for pain…	…on practice registration and annually.
Development	MacLean[127,128]	2004	Function assessment	IF a patient is diagnosed with symptomatic osteoarthritis of the knee or hip, THEN his or her functional status should be assessed annually and when new to a practice	Patients with symptomatic OA…	…should be assessed for function…	…on practice registration and annually.
Development	Steel[124]	2004	Pain & function assessment	IF a person aged 65 or older is treated for symptomatic osteoarthritis, THEN functional status and degree of pain should be assessed at least annually	Patients with symptomatic OA…	…should be assessed for pain and function…	…annually.
Development	Saliba[133]	2005	Examination	IF a NH resident has a new joint pain that is reported to the primary care provider, THEN the affected joint and periarticular structures should be examined within 1 month or there should be documentation that the problem has resolved.	Patients with joint pain…	…should have documentary evidence of a joint examination…	…within four weeks of identification.
Development	Saliba[133]	2005	Examination	IF a non-OTC drug is newly prescribed to treat new joint pain THEN evidence that the affected joint was examined should be documented within 4 weeks.	Patients with joint pain who receive treatment…	…should have documentary evidence of a joint examination…	…within four weeks of prescription of a new drug.
Development	Saliba[133]	2005	Differential diagnosis assessment	IF an NH resident has monoarticular joint pain associated with redness, warmth, and/or swelling, and the patient also has an oral temperature ›38^{o}C, and does not have a previously established diagnosis of gout or pseudogout, THEN a diagnostic aspiration of the painfully swollen red joint should be performed that day	Patients with a possible differential diagnosis of a septic joint…	…should have a joint aspiration…	…the same day.

Development	PCPI[129]	2006	Pain & function assessment	Percentage of patient visits for patients aged 21 years and older with a diagnosis of osteoarthritis with assessment for function and pain	Patients with a diagnosis of OA…	…should be assessed for pain and function.	-
Development	PCPI[129]	2006	Examination	Percentage of patients aged 21 years and older with a diagnosis of OA for whom a physical examination of the involved joint was performed during the initial visit	Patients with joint pain…	…should have documentary evidence of a joint examination…	…at the initial assessment.
Development	MacLean[132]	2007	Pain assessment	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 [BECAUSE this information should direct therapeutic decisions].	Patients with symptomatic OA (hip or knee)…	…should be assessed for pain…	…on practice registration and annually.
Development	MacLean[132]	2007	Function assessment	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, BECAUSE this information should direct therapeutic decisions.	Patients with symptomatic OA (hip or knee)…	…should be assessed for function…	…on practice registration and annually.
Development	Smith[123]	2007	Pain & function assessment	IF a homebound patient is diagnosed with symptomatic osteoarthritis, THEN his or her functional status and the degree of pain should be assessed at each visit.	Patients with symptomatic OA…	…should be assessed for pain…	…at each visit.
Development	Smith[123]	2007	Differential diagnosis assessment	IF a homebound patient has monoarticular joint pain associated with redness, warmth, or swelling AND the patient also has an oral temperature greater than 38.0 °C and does not have a previously established diagnosis of pseudogout or gout, THEN diagnostic aspiration of the painfully swollen, red joint should be performed that day.	Patients with a possible differential diagnosis of a septic joint…	…should have a joint aspiration…	…the same day.

Development	EUMUSC.net[125]	2012	Pain & function assessment	If a patient is diagnosed with symptomatic osteoarthritis (OA), then he or she should be assessed for pain, functional ability, level of physical activity, body mass index (BMI), and labor force participation at baseline or when experiencing significant pain or functional limitation	Patients with symptomatic OA…	…should be assessed for pain, function, physical activity, BMI, and work capability…	…at diagnosis and each exacerbation.
Implementation	Wenger[78]	2003	Pain & function assessment	IF a vulnerable elder receives a diagnosis of symptomatic osteoarthritis, THEN functional status and degree of pain should be assessed annually.	Patients with symptomatic OA should be assessed annually for pain and function		
Implementation	Wenger[78]	2003	Differential diagnosis assessment	IF a vulnerable elder has monoarticular joint pain associated with redness, warmth, or swelling and the patient also has an oral temperature greater than 38.0 °C and does not have a previously established diagnosis of pseudogout or gout, THEN a diagnostic aspiration of the painfully swollen red joint should be performed that day.	Patients with a possible differential diagnosis of a septic joint…	…should have a joint aspiration…	…the same day.
Implementation	McGlynn[77]	2003	Pain & function assessment	Providers caring for patients with symptoms of OA should document all at least one of the following at least once in 2 years: the location of symptoms, and/or the presence or absence of limitations in daily activities	Patients with symptomatic OA…	…should be assessed for symptoms and function…	…every two years.
Implementation	Chodosh[73]	2004	Pain & function assessment	IF a vulnerable elder is diagnosed with symptomatic osteoarthritis, THEN functional status and degree of pain should be assessed annually.	Patients with symptomatic OA…	…should be assessed for pain and function…	…annually.

Implementation	Asch[71]	2004	Pain & function assessment	Providers caring for patients with symptoms of osteoarthritis should document all of the following at least once in 2 years: the location of symptoms and/or the presence or absence of limitations in daily activities	Patients with symptomatic OA…	…should be assessed for symptoms and function…	…every two years.
Implementation	Ganz[74]	2006	Pain & function assessment	IF a person age 75 or older is diagnosed with symptomatic osteoarthritis, THEN functional status and degree of pain should be assessed annually.	Patients with symptomatic OA…	…should be assessed for pain and function…	…annually.
Implementation	Steel[68]	2007	Pain & function assessment	The percentage of patients treated for symptomatic osteoarthritis, whose notes contain a record that they have been assessed for (1) functional status and (2) degree of pain in the last year	Patients with symptomatic OA…	…should be assessed for pain and function…	…annually.
Implementation	Broadbent[69]	2008	Function assessment	The percentage of patients treated for symptomatic osteoarthritis, whose notes contain a record that they have been assessed for functional status in the last year	Patients with symptomatic OA…	…should be assessed for function…	…annually.
Implementation	Broadbent[69]	2008	Pain assessment	The percentage of patients treated for symptomatic osteoarthritis, whose notes contain a record that they have been assessed for degree of pain in the last year	Patients with symptomatic OA…	…should be assessed for pain …	…annually.
Implementation	Østerås[82]	2013	Pain assessment	If you have pain, has it been assessed in the past year?	Patients with symptomatic OA…	…should be assessed for pain …	…annually.
Implementation	Østerås[82]	2013	Pain assessment	If you have had problems related to daily activities, have these problems been assessed by health personnel in the past year?	Patients with symptomatic OA…	…should be assessed for function…	…annually.

	Synthesised indicators		
	Patients with symptomatic OA…	…should be assessed for pain …	…annually.
	Patients with symptomatic OA…	…should be assessed for function…	…annually.
	Patients with joint pain…	*…should have documentary evidence of a joint examination.*	*[not selected for inclusion]*
	Patients with a possible differential diagnosis of a septic joint…	*…should have a joint aspiration…*	*…the same day.* *[not selected for inclusion]*

Appendix E. Template development studies appendices (Chapter Three)

Appendix E.1. Read codes for assessment of routinely recorded management of OA

Assessment of pain (Pain assessment, not pain location, codes):	
1D13. C/O: a pain	1M5.. Pattern of pain (+ child codes)
1DC.. Pain character (+ child codes)	2IA.. O/E - painful sign (+ child codes)
1M2.. Pain score	
Assessment of function	
Function assessment codes:	39A.. Bathing ability (+ child codes)
13C.. Mobility - social functioning (+ child codes)	39B.. Walking aid use (+ child codes)
1DD.. Symptom restricts activity	39D.. Walking distance (+ child codes)
1P8.. Ability to perform personal care activity (+ child codes)	39E.. Physical disabil.assess score (+ child codes)
1PA.. Ability to perform activities of everyday life (+ child codes)	39F.. Barthel index
39... Disability assessment-physical	39G.. Needs help with housework (+ child codes)
392.. Grooming ability (+ child codes)	39J.. Independent in cooking
395.. Dressing ability (+ child codes)	39K.. Independent in housework
396.. Chair/bed transfer (+ child codes)	68O.. Mobility screen
397.. Toilet dependency (+ child codes)	ZV4L0 [V] Poor mobility
399.. Stairs – ability (+ child codes)	ZV4KA [V]Limitation of activities due to disability
Evidence of education or advice	
679M. Health education – rheumatology	679Z. Health education - subject NOS
Exercise assessment	
138.. Exercise grading (+ child codes)	8E7.. Physical exercises (+ child codes)
Exercise advice	
6798. Health ed. – exercise	8CA5. Patient advised re exercise (+ child codes)
Weight loss advice for people with a BMI ≥25kgm^{-2}	
38F.. Assessment for bariatric surgery	8CV7. Anti-obesity drug therapy commenced
66C.. Obesity monitoring (+ child codes)	8H4n. Referral to weight management special interest general practitioner
66CH. Weight management plan started	8H5N. Referral to bariatric surgeon
67I9. Advice about weight	8H76. Refer to dietician (+ child codes)
8CA4. Patient advised re diet	8HHH. Refer to weight management programme (+ child codes)
8CA40 Patient advised to lose weight	8Hlv. Referral for pre-bariatric surgery assessment
8Cd7. Advice given about weight management	8IAM. Referral to weight management service declined
8CP4. Discussion about bariatric operative procedure	8IAu. Weight management advice declined
8CP5. Discussion about weight management programme	9N1yF Seen in bariatric surgery clinic
8CT5. Anti-obesity drug therapy discontinued	9NS03 Referral to weight management service offered
Evidence of documentation of NSAID risk assessment	
8CD.. Usual warning given	9c0K. Risk information
9OhB. Non-steroidal anti-inflammatory drug risk assessment completed	

Appendix E.2. Initial draft template

Assessment

Read code	Read term	Domain
1M2	Pain score	Pain
7L1Wz	Assessment NOS	Function
13C.. (+child codes)	Mobility - social functioning	Function
22A.. (+child codes)	O/E – weight	Weight (for BMI)
229.. (+child codes)	O/E - height	Height (for BMI)
22K.. (+child codes)	Body Mass Index	BMI
C380. (+child codes)	Obesity	BMI
138.. (+child codes)	Exercise grading	Exercise assessment
68L.. (+child codes)	Exercise status screening	Exercise assessment
8I2T.	Non-steroidal anti-inflammatory drugs contraindicated	NSAID risk assessment
TJ56.	Adverse reaction to non-steroidal anti-inflammatory drugs	NSAID risk assessment
U6053	[X]Other non-steroidal anti-inflammatory drugs [NSAID] causing adverse effects in therapeutic use	NSAID risk assessment
Unclear how to include OTC paracetamol/NSAID code – no obvious code		

Clinical management

Read code	Read term	Domain
6799.	Health ed. – diet	Weight advice
679P.	Health education - weight management (or system default health promotion code)	Weight advice
67I9.	Advice about weight	Weight advice
66C..	Weight monitoring	Weight advice
8B57.	Weight reducing diet	Weight advice
8CA40	Patient advised to lose weight	Weight advice
6798.	Health ed. – exercise or system default	Exercise advice
679..	Health education – subject (need to specify in free text)	
679M.	Health education - rheumatology (for the guidebook?) [? use a synonymous created term)	
9D1..	MED3 - doctor's statement	
9D2..	MED5 - doctor's special stat.	
	Referral (link to new template)	

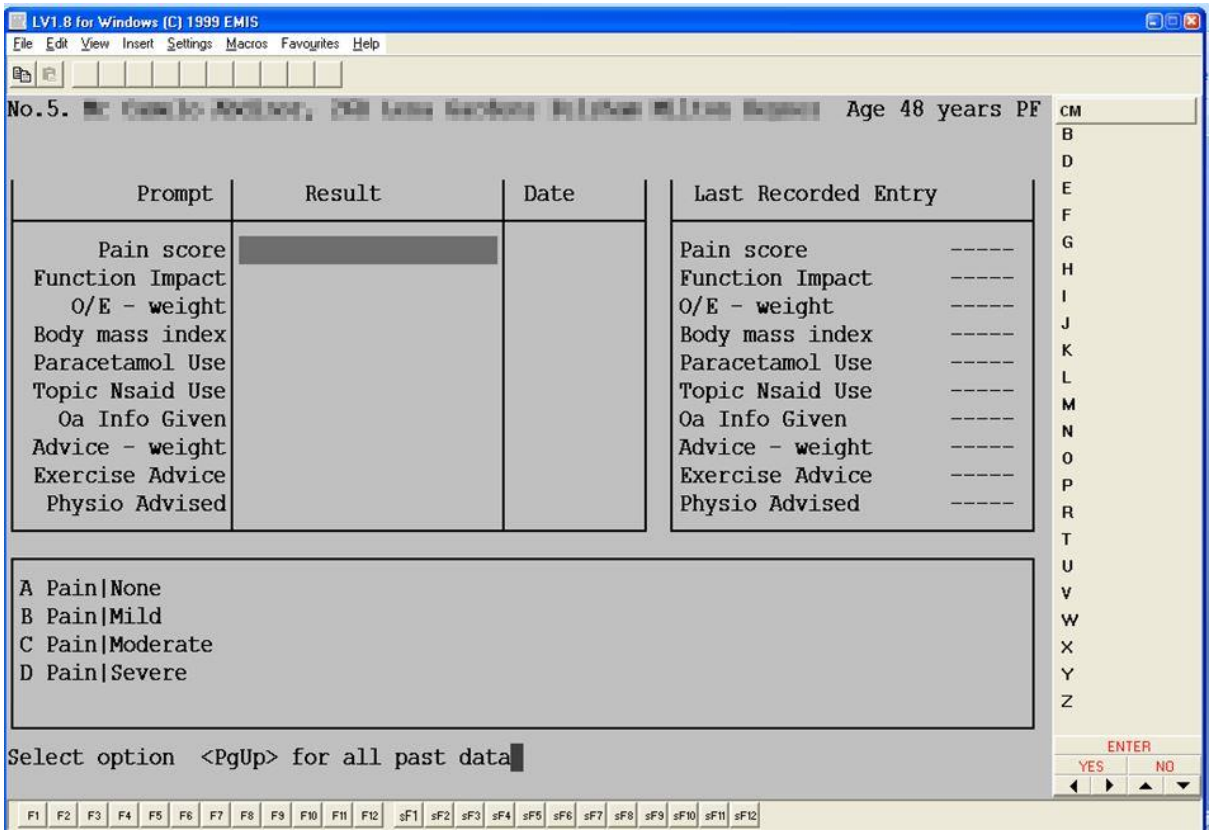
Referrals should link to another template, below.

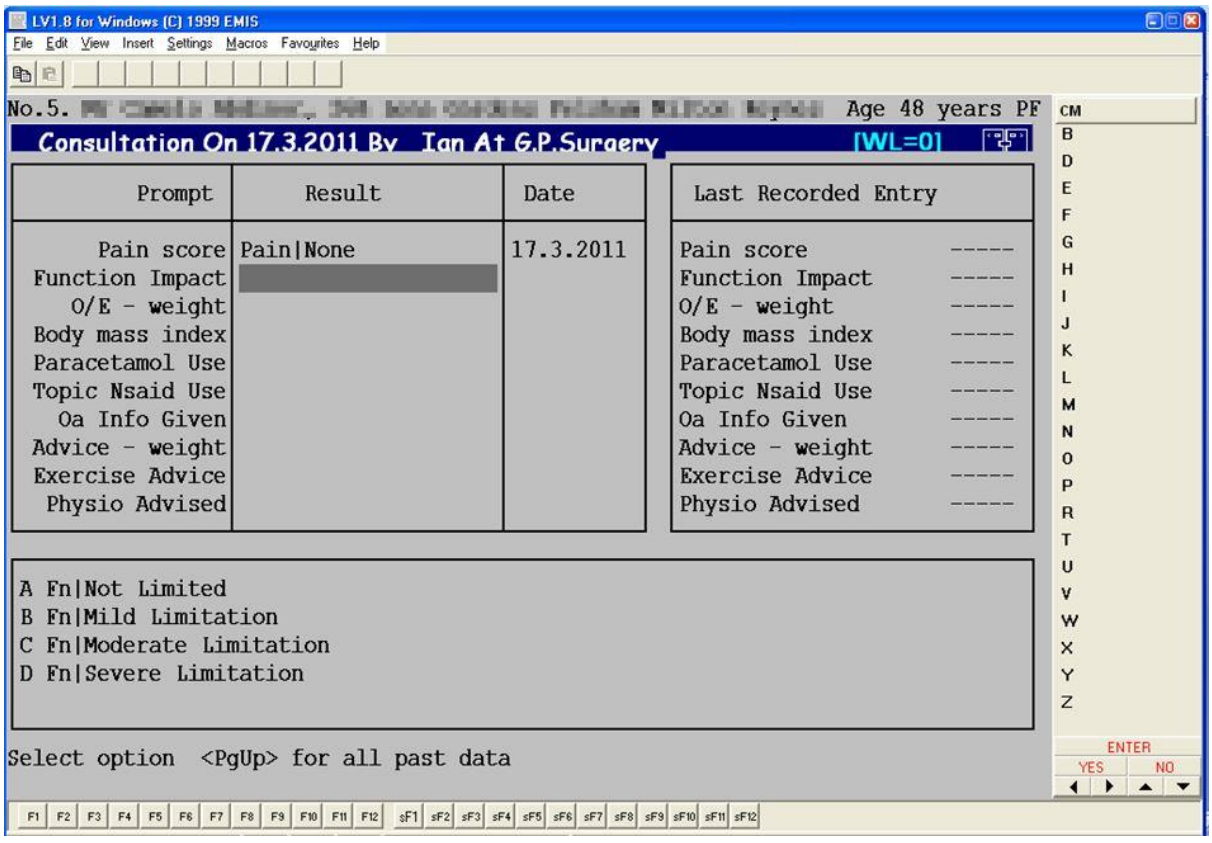
[Prescribing information to be collected for additional information outside of the template]

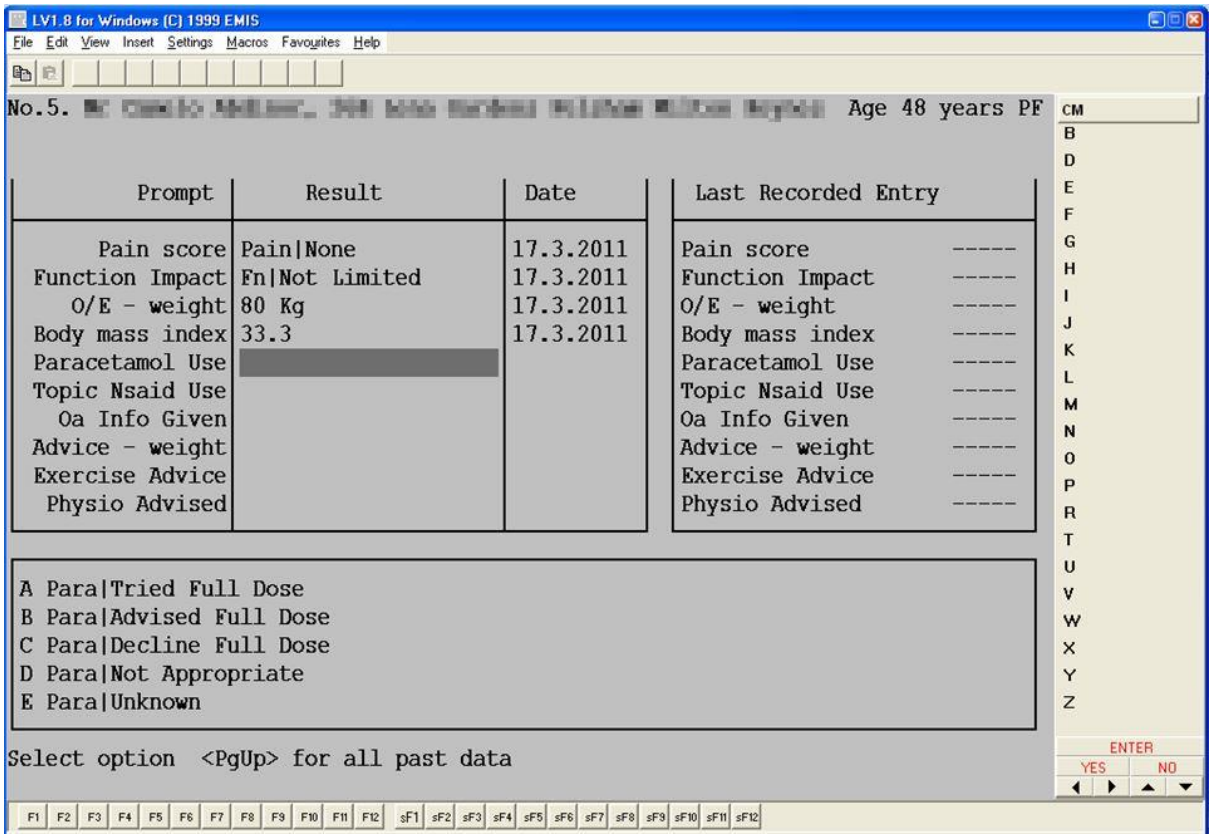
Referrals

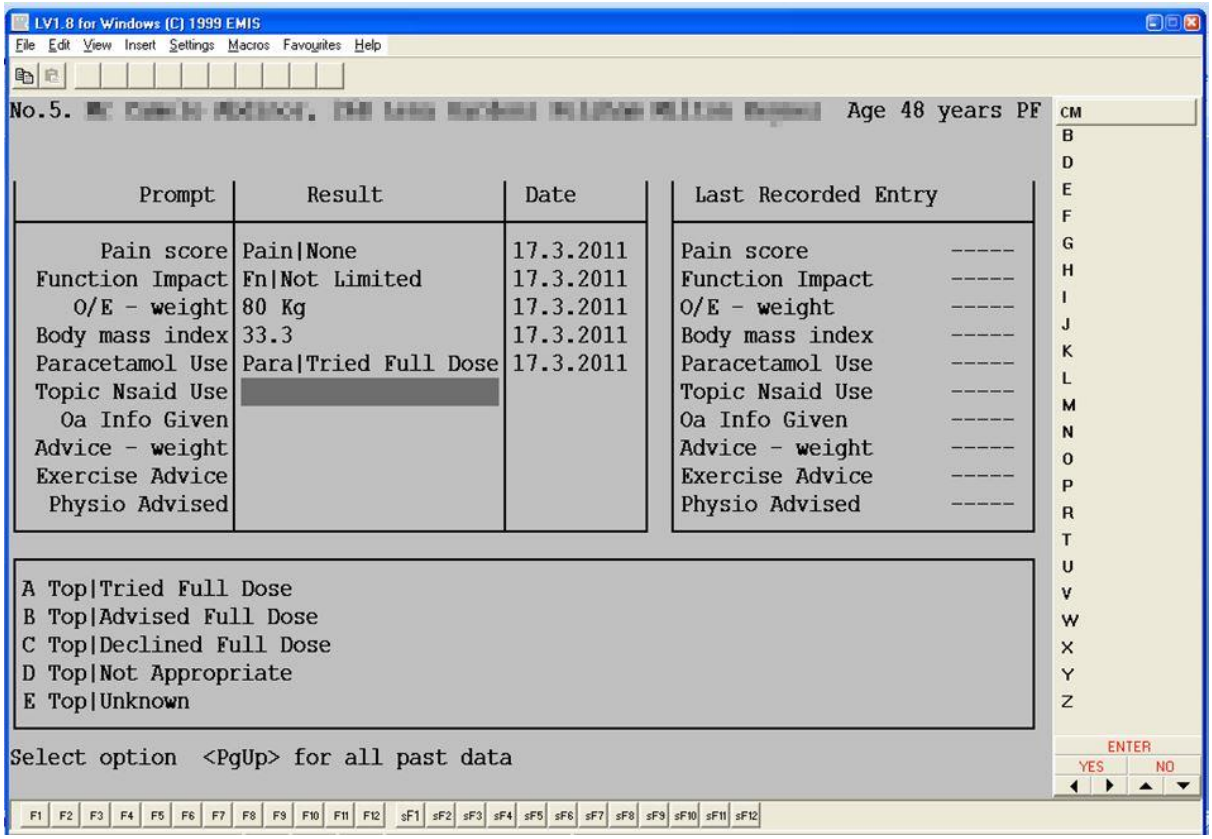
Read code	Read term	Domain
8BAH.	Exercise on prescription	Exercise referral/advice
8CAc.	Advised to contact physiotherapy triage service	Exercise referral/advice
8E... (+child codes)	Physiotherapy/remedial therapy	Exercise referral/advice
8H77.	Refer to physiotherapist	Exercise referral/advice
8H7q.	Referral for exercise therapy	Exercise referral/advice
8H7s.	Referral to physical activity programme	Exercise referral/advice
8HHc.	Referred for exercise programme	Exercise referral/advice
9NJ3.	In-house physio	Exercise referral/advice
9NJ4.	In-house physiotherapy - domiciliary visit	Exercise referral/advice
9NJk.	In-house physiotherapy first appointment	Exercise referral/advice
9NJm.	In-house physiotherapy follow-up appointment	Exercise referral/advice
9N0F.	Seen in physiotherapy dept	Exercise referral/advice
8H76.	Refer to dietician	Weight management referral
8HHH.	Refer to weight management programme	Weight management referral
8H7J.	Refer to occupational therap.	OT referral (non-ambulatory aids/devices)
8H7Q.	Refer to surgical fitter	Podiatry referral (ambulatory aids/devices)
8H7R.	Refer to chiropodist	Podiatry referral (ambulatory aids/devices)
8H7S.	Refer to orthotist	Podiatry referral (ambulatory aids/devices)
8H7k.	Referral to community-based podiatry service	Podiatry referral (ambulatory aids/devices)
8H7l.	Referral to hospital-based podiatry service	Podiatry referral (ambulatory aids/devices)
8H7m.	Referral to private state registered podiatry service	Podiatry referral (ambulatory aids/devices)
8H4B.	Referred to rheumatologist	Secondary care referral
8H54.	Orthopaedic referral	Secondary care referral
8H69.	Refer to pain clinic	Secondary care referral

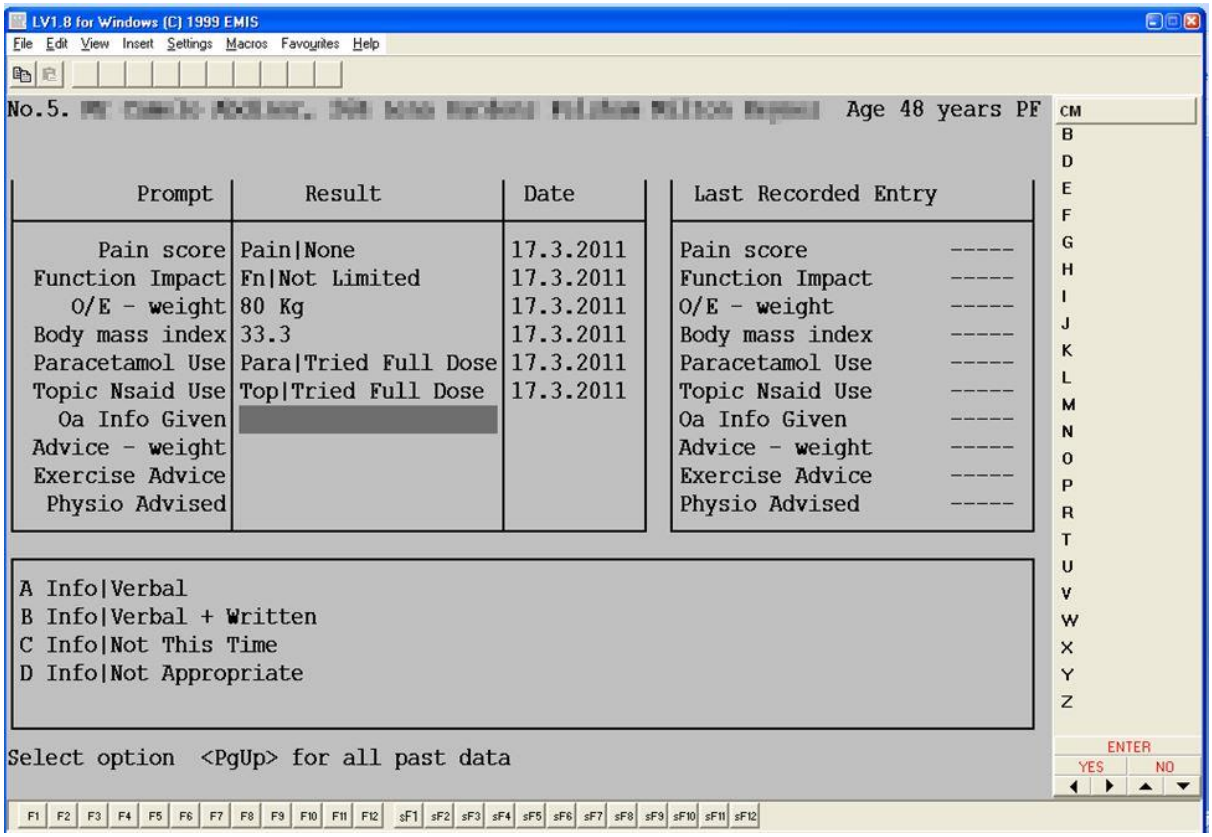
Appendix E.3. Final template created by the CRN staff to the candidate's specification

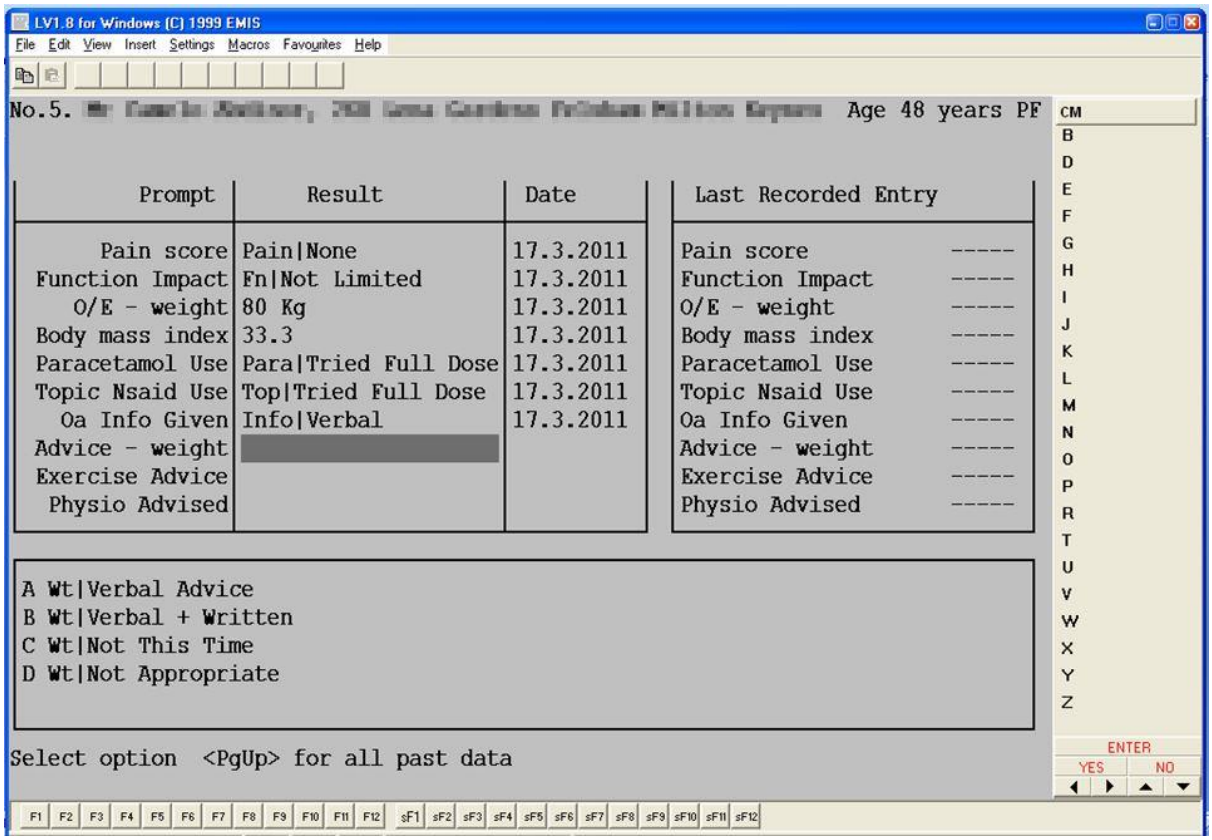


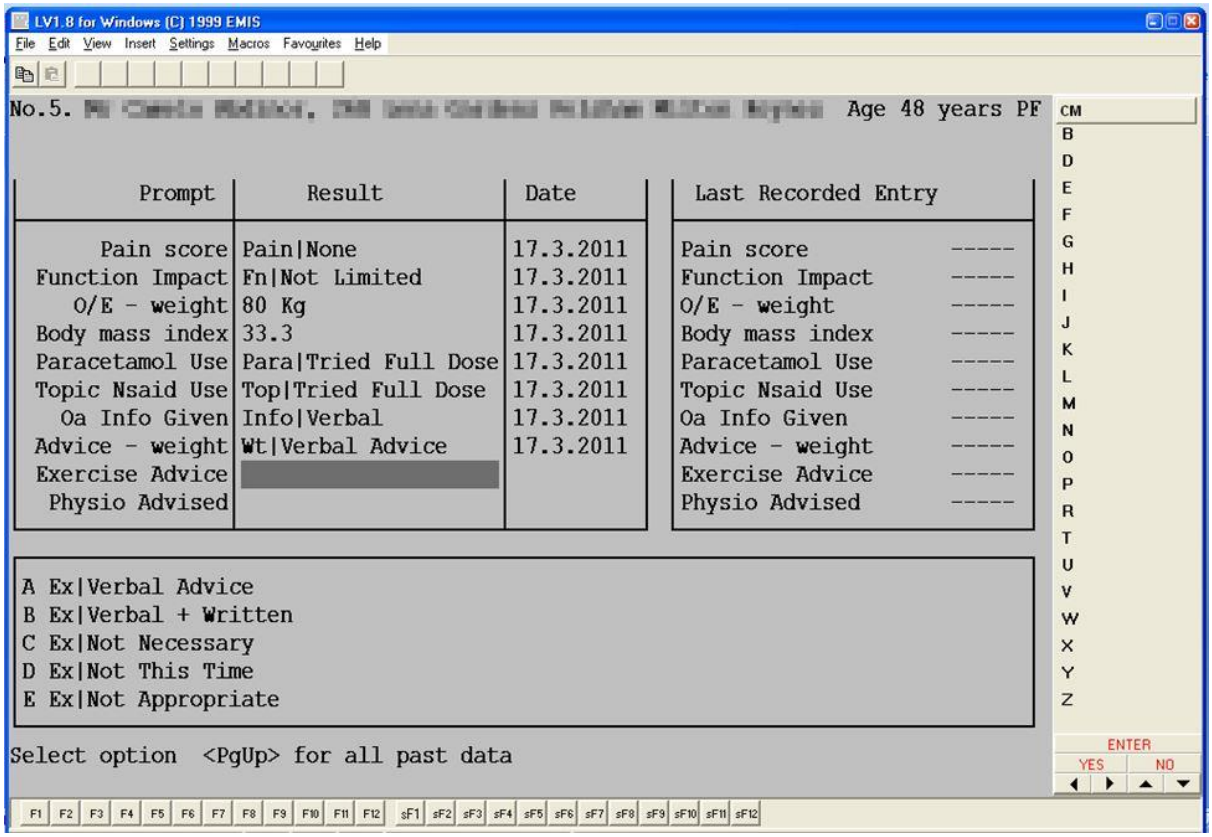


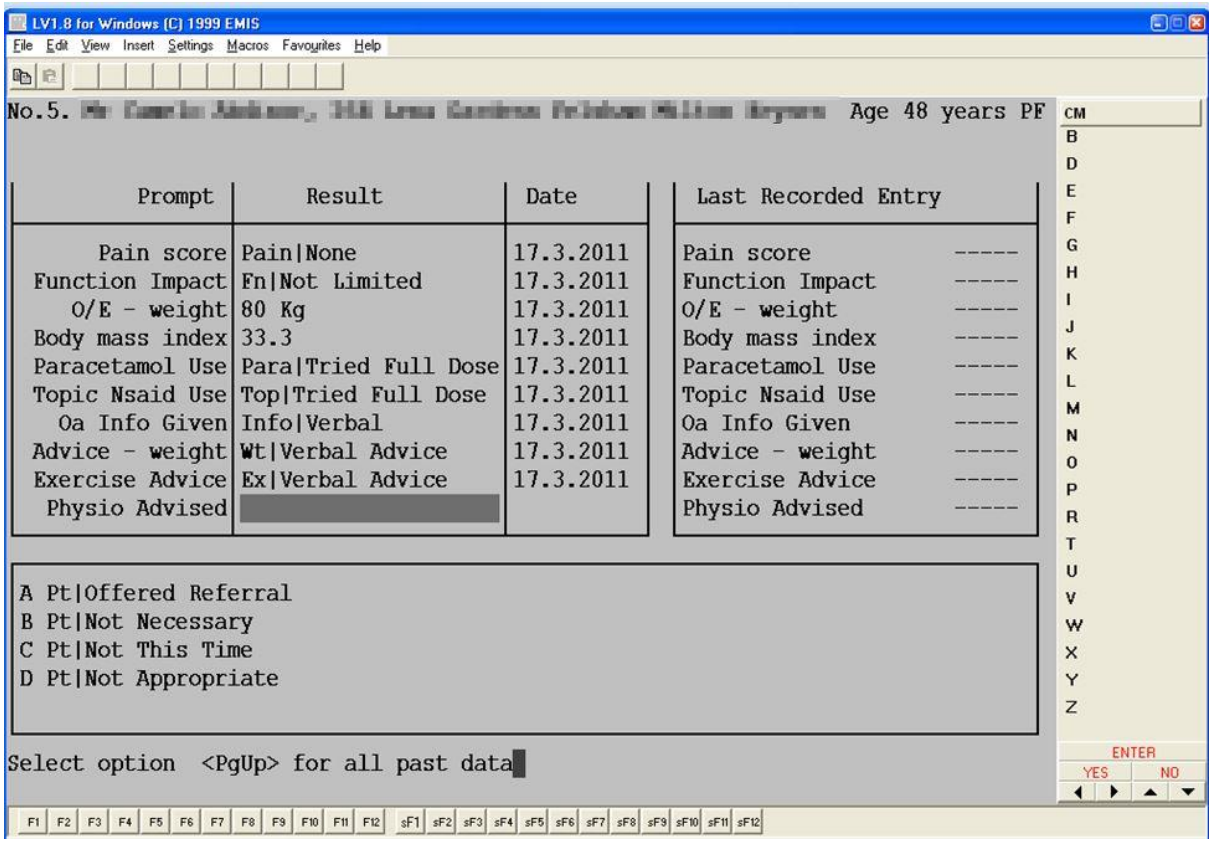

LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
CM
Prompt | Result | Date | Last Recorded Entry
Pain score | Pain|None | 17.3.2011 | Pain score -----
Function Impact | Fn|Not Limited | 17.3.2011 | Function Impact -----
O/E - weight | 80 Kg | 17.3.2011 | O/E - weight -----
Body mass index | 33.3 | 17.3.2011 | Body mass index -----
Paracetamol Use | Paracetamol Use -----
Topic Nsaid Use | Topic Nsaid Use -----
Oa Info Given | Oa Info Given -----
Advice - weight | Advice - weight -----
Exercise Advice | Exercise Advice -----
Physio Advised | Physio Advised -----
A Para|Tried Full Dose
B Para|Advised Full Dose
C Para|Decline Full Dose
D Para|Not Appropriate
E Para|Unknown
Select option <PgUp> for all past data
B D E F G H I J K L M N O P R T U V W X Y Z
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12

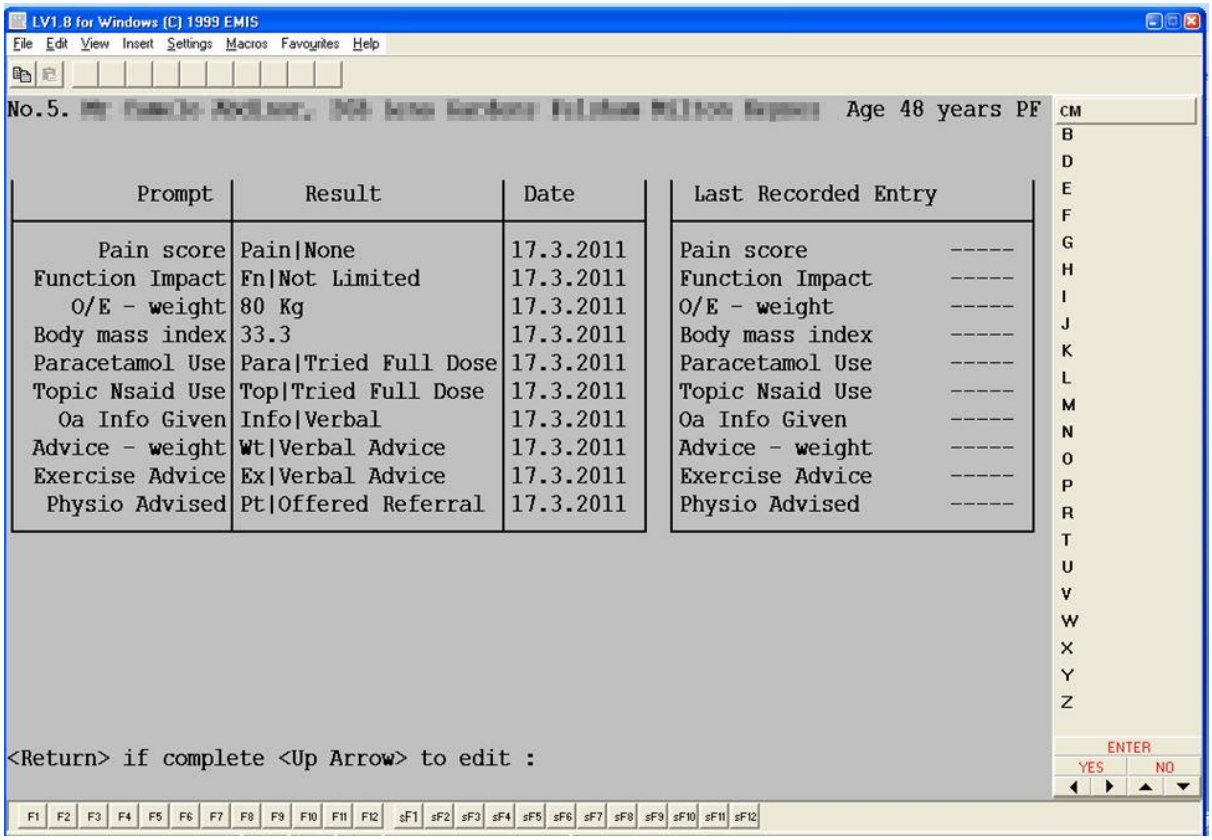

LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
CM
Prompt | Result | Date | Last Recorded Entry
Pain score | Pain|None | 17.3.2011 | Pain score -----
Function Impact | Fn|Not Limited | 17.3.2011 | Function Impact -----
O/E - weight | 80 Kg | 17.3.2011 | O/E - weight -----
Body mass index | 33.3 | 17.3.2011 | Body mass index -----
Paracetamol Use | Para|Tried Full Dose | 17.3.2011 | Paracetamol Use -----
Topic Nsaid Use | Topic Nsaid Use -----
Oa Info Given | Oa Info Given -----
Advice - weight | Advice - weight -----
Exercise Advice | Exercise Advice -----
Physio Advised | Physio Advised -----
A Top|Tried Full Dose
B Top|Advised Full Dose
C Top|Declined Full Dose
D Top|Not Appropriate
E Top|Unknown
Select option <PgUp> for all past data
B D E F G H I J K L M N O P R T U V W X Y Z
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12

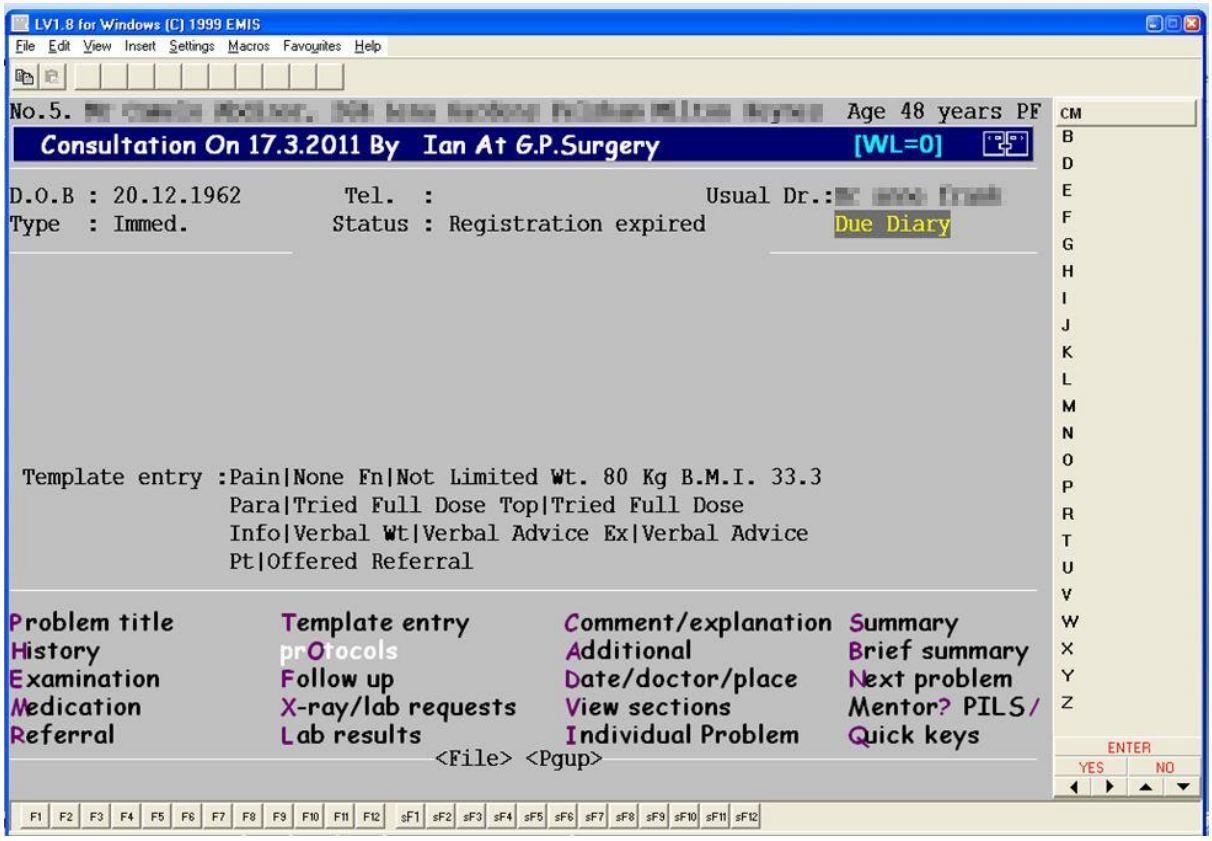

LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
CM
Prompt | Result | Date
Pain score | Pain|None | 17.3.2011
Function Impact | Fn|Not Limited | 17.3.2011
O/E - weight | 80 Kg | 17.3.2011
Body mass index | 33.3 | 17.3.2011
Paracetamol Use | Para|Tried Full Dose | 17.3.2011
Topic Nsaid Use | Top|Tried Full Dose | 17.3.2011
Oa Info Given
Advice - weight
Exercise Advice
Physio Advised
Last Recorded Entry
Pain score -----
Function Impact -----
O/E - weight -----
Body mass index -----
Paracetamol Use -----
Topic Nsaid Use -----
Oa Info Given -----
Advice - weight -----
Exercise Advice -----
Physio Advised -----
A Info|Verbal
B Info|Verbal + Written
C Info|Not This Time
D Info|Not Appropriate
Select option <PgUp> for all past data
B D E F G H I J K L M N O P R T U V W X Y Z
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12


LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
CM
Prompt | Result | Date
Pain score | Pain|None | 17.3.2011
Function Impact | Fn|Not Limited | 17.3.2011
O/E - weight | 80 Kg | 17.3.2011
Body mass index | 33.3 | 17.3.2011
Paracetamol Use | Para|Tried Full Dose | 17.3.2011
Topic Nsaid Use | Top|Tried Full Dose | 17.3.2011
Oa Info Given | Info|Verbal | 17.3.2011
Advice - weight
Exercise Advice
Physio Advised
Last Recorded Entry
Pain score -----
Function Impact -----
O/E - weight -----
Body mass index -----
Paracetamol Use -----
Topic Nsaid Use -----
Oa Info Given -----
Advice - weight -----
Exercise Advice -----
Physio Advised -----
A Wt|Verbal Advice
B Wt|Verbal + Written
C Wt|Not This Time
D Wt|Not Appropriate
Select option <PgUp> for all past data
B D E F G H I J K L M N O P R T U V W X Y Z
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12


LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
CM
Prompt | Result | Date | Last Recorded Entry
Pain score | Pain|None | 17.3.2011 | Pain score -----
Function Impact | Fn|Not Limited | 17.3.2011 | Function Impact -----
O/E - weight | 80 Kg | 17.3.2011 | O/E - weight -----
Body mass index | 33.3 | 17.3.2011 | Body mass index -----
Paracetamol Use | Para|Tried Full Dose | 17.3.2011 | Paracetamol Use -----
Topic Nsaid Use | Top|Tried Full Dose | 17.3.2011 | Topic Nsaid Use -----
Oa Info Given | Info|Verbal | 17.3.2011 | Oa Info Given -----
Advice - weight | Wt|Verbal Advice | 17.3.2011 | Advice - weight -----
Exercise Advice | | | Exercise Advice -----
Physio Advised | | | Physio Advised -----
A Ex|Verbal Advice
B Ex|Verbal + Written
C Ex|Not Necessary
D Ex|Not This Time
E Ex|Not Appropriate
Select option <PgUp> for all past data
B D E F G H I J K L M N O P R T U V W X Y Z
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12


LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
CM
Prompt | Result | Date | Last Recorded Entry
Pain score | Pain|None | 17.3.2011 | Pain score -----
Function Impact | Fn|Not Limited | 17.3.2011 | Function Impact -----
O/E - weight | 80 Kg | 17.3.2011 | O/E - weight -----
Body mass index | 33.3 | 17.3.2011 | Body mass index -----
Paracetamol Use | Para|Tried Full Dose | 17.3.2011 | Paracetamol Use -----
Topic Nsaid Use | Top|Tried Full Dose | 17.3.2011 | Topic Nsaid Use -----
Oa Info Given | Info|Verbal | 17.3.2011 | Oa Info Given -----
Advice - weight | Wt|Verbal Advice | 17.3.2011 | Advice - weight -----
Exercise Advice | Ex|Verbal Advice | 17.3.2011 | Exercise Advice -----
Physio Advised | | | Physio Advised -----
A Pt|Offered Referral
B Pt|Not Necessary
C Pt|Not This Time
D Pt|Not Appropriate
Select option <PgUp> for all past data
B D E F G H I J K L M N O P R T U V W X Y Z
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12


LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
Prompt | Result | Date
Pain score | Pain|None | 17.3.2011
Function Impact | Fn|Not Limited | 17.3.2011
O/E - weight | 80 Kg | 17.3.2011
Body mass index | 33.3 | 17.3.2011
Paracetamol Use | Para|Tried Full Dose | 17.3.2011
Topic Nsaid Use | Top|Tried Full Dose | 17.3.2011
Oa Info Given | Info|Verbal | 17.3.2011
Advice - weight | Wt|Verbal Advice | 17.3.2011
Exercise Advice | Ex|Verbal Advice | 17.3.2011
Physio Advised | Pt|Offered Referral | 17.3.2011
Last Recorded Entry
Pain score -----
Function Impact -----
O/E - weight -----
Body mass index -----
Paracetamol Use -----
Topic Nsaid Use -----
Oa Info Given -----
Advice - weight -----
Exercise Advice -----
Physio Advised -----
CM
B D E F G H I J K L M N O P R T U V W X Y Z
<Return> if complete <Up Arrow> to edit :
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12


LV1.8 for Windows (C) 1999 EMIS
File Edit View Insert Settings Macros Favourites Help
No.5. [redacted] Age 48 years PF
Consultation On 17.3.2011 By Ian At G.P.Surgery [WL=0]
D.O.B : 20.12.1962 Tel. : Usual Dr.: [redacted]
Type : Immed. Status : Registration expired Due Diary
Template entry :Pain|None Fn|Not Limited Wt. 80 Kg B.M.I. 33.3
Para|Tried Full Dose Top|Tried Full Dose
Info|Verbal Wt|Verbal Advice Ex|Verbal Advice
Pt|Offered Referral
Problem title Template entry Comment/explanation Summary
History prOtocols Additional Brief summary
Examination Follow up Date/doctor/place Next problem
Medication X-ray/lab requests View sections Mentor? PILS/
Referral Lab results Individual Problem Quick keys
<File> <Pgup>
CM
B D E F G H I J K L M N O P R T U V W X Y Z
ENTER
YES NO
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12 sF1 sF2 sF3 sF4 sF5 sF6 sF7 sF8 sF9 sF10 sF11 sF12

Appendix E.4. Draft MOSAICS OA definition codes with GP consensus exercise results and CiPCA testing outcomes

Read code	Separately-considered child codes	Number of relevant child codes	Read Term	GP1	GP2	GP3	GP4	GP5	GP6	Outcome
1M0..		1	Pain in upper limb	Include	Exclude	Exclude	Include	Include	Exclude	Include
	1M00.		Pain in elbow	Include	Exclude	Exclude	Include	Include	Exclude	Include
1M1..		4	Pain in lower limb	Include	Exclude	Exclude	Include	Include	Exclude	Include
	1M10.		Knee pain	Include	Exclude	Include	Include	Include	Include	Include
	1M11.		Foot pain	Include	Exclude	Include	Include	Include	Possible	Include
	1M12.		Anterior knee pain	Include	Exclude	Include	Include	Include	Exclude	Include
	1M13.		Ankle pain	Include	Exclude	Unsure	Include	Include	Possible	Include
2G26.	No		O/E - hands - Heberden's nodes	Include	Include	Include	Include	Include	Include	Include
N05..	(grouped)	98	Osteoarthritis and allied disorders	Include	Include	Include	Include	Include	Include	Include
N064.	(grouped)	31	Transient arthropathy	Exclude	Exclude	Include	Exclude	Exclude	Exclude	Exclude
N065.		12	Unspecified polyarthropathy or polyarthritis	Exclude	Possibly	Include	Include	Exclude	Probably	Include
	N065A		Generalised arthritis	Include	Include	Include	Include	Include	Include	Include
N066.	(grouped)	10	Unspecified monoarthritis [synonym = coxitis]	Exclude	Possibly	Include	Include	Exclude	Include	Include
N06y.	(grouped)	11	Other specified arthropathy	Exclude	Possibly	Include	Unsure	Exclude	Exclude	Exclude
N06z.	(grouped)	13	Arthropathy NOS [synonym = arthritis]	Include	Possibly	Include	Include	Exclude	Include	Include
N094.	(grouped)	32	Pain in joint – arthralgia	Include	Exclude	Include	Include	Include	Possible	Include
N095.	(grouped)	32	Joint stiffness NEC	Exclude	Exclude	Include	Include	Undecided	Exclude	Exclude
N096.	(grouped)	31	Other joint symptoms [synonyms = joint crepitus, musculoskeletal pain - joints]	Include	Exclude	Include	Include	Include	Probably	Include
N097.	(grouped)	8	Difficulty in walking	Exclude	Exclude	Include	Exclude	Include	Exclude	Exclude
N099.	(grouped)	22	Clicking joint	Exclude	Exclude	Include	Exclude	Exclude	Exclude	Exclude

Read code	Separately-considered child codes	Number of relevant child codes	Read Term	GP1	GP2	GP3	GP4	GP5	GP6	Outcome
N09A.		1	Patellofemoral disorder	Include	Exclude	Exclude	Unsure	Unsure	Exclude	Exclude
	N09AX		Disorder of patella, unspecified	Include	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
N09y.	(grouped)	11	Other specified joint disorders	Exclude	Exclude	Exclude	Exclude	Exclude	Probably	Exclude
N09z.	(grouped)	11	Joint disorders NOS	Exclude	Exclude	Exclude	Exclude	Exclude	Possibly	Exclude
N240.	(grouped)	8	Rheumatism and fibrositis unspecified	Exclude	Exclude	Unsure	Exclude	Exclude	Possibly	Exclude
N241.	(grouped)	5	Myalgia and myositis unspecified	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
N245.		8	Pain in limb	Include	Exclude	Exclude	Include	Include	Exclude	Include
	N2450		Hand pain	Include	Exclude	Unsure	Include	Include	Exclude	Include
	N2451		Foot pain	Include	Exclude	Unsure	Include	Include	Exclude	Include
	N2452		Pain in leg	Include	Exclude	Exclude	Exclude	Include	Exclude	Exclude
	N2453		Pain in arm	Include	Exclude	Exclude	Exclude	Include	Exclude	Exclude
	N2454		Calf pain	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
	N2455		Axillary pain	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
	N2456		Tender heel pad	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
	N2457		Shoulder pain	Include	Exclude	Exclude	Exclude	Include	Exclude	Exclude
N247.		4	Other musculoskeletal limb symptoms	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
	N2470		Swelling of limb	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
	N2471		Leg cramps	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
	N2472		Cramp	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
	N247z		Musculoskeletal limb symptoms NOS	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
R01z2	No		[D]Musculoskeletal pain	Include	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude
R01zz	No		[D]Nervous or musculoskeletal symptoms NOS	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude	Exclude

Appendix E.5. CiPCA assessment of Read code triggers for the MOSAICS template

CODE	TERM	CODE	TERM
1M10	**Knee pain**	*N051A*	*Coxarthrosis resulting from dysplasia, bilateral*
1M11	**Foot pain**	*N051B*	*Primary gonarthrosis, bilateral*
1M13	**Ankle pain**	*N051E*	*Localised, primary osteoarthritis of toe*
N05	**Osteoarthritis and allied disorders**	*N051F*	*Localised, primary osteoarthritis of elbow*
N050	**Generalised osteoarthritis - OA**	*N051z*	*Localised, primary osteoarthritis NOS*
N0500	*Generalised osteoarthritis of unspecified site*	*N052*	*Localised, secondary osteoarthritis*
N0501	**Generalised osteoarthritis of the hand**	*N0520*	*Localised, secondary osteoarthritis of unspecified site*
N0502	**Generalised osteoarthritis of multiple sites**	*N0521*	*Localised, secondary osteoarthritis of the shoulder region*
N0503	*Bouchard's nodes with arthropathy*	*N0522*	*Localised, secondary osteoarthritis of the upper arm*
N0504	**Primary generalized osteoarthrosis**	*N0523*	*Localised, secondary osteoarthritis of the forearm*
N0505	**Secondary multiple arthrosis**	*N0524*	*Localised, secondary osteoarthritis of the hand*
N0506	*Erosive osteoarthrosis*	*N0525*	*Localised, secondary osteoarthritis of the pelvic region and thigh*
N0507	*Heberden's nodes with arthropathy*	*N0526*	*Localised, secondary osteoarthritis of the lower leg*
N050z	**Generalised osteoarthritis NOS**	*N0527*	*Localised, secondary osteoarthritis of the ankle and foot*
N051	*Localised, primary osteoarthritis*	*N0528*	*Localised, secondary osteoarthritis of other specified site*
N0510	*Localised, primary osteoarthritis of unspecified site*	*N0529*	*Post-traumatic coxarthrosis, bilateral*
N0511	*Localised, primary osteoarthritis of the shoulder region*	*N052A*	*Post-traumatic gonarthrosis, bilateral*
N0512	*Localised, primary osteoarthritis of the upper arm*	*N052B*	*Post-traumatic arthrosis of first carpometacarpal joints, bilateral*
N0513	*Localised, primary osteoarthritis of the forearm*	*N052C*	*Post-traumatic gonarthrosis, unilateral*
N0514	**Localised, primary osteoarthritis of the hand**	*N052z*	*Localised, secondary osteoarthritis NOS*
N0515	*Localised, primary osteoarthritis of the pelvic region and thigh*	*N053*	*Localised osteoarthritis, unspecified*
N0516	*Localised, primary osteoarthritis of the lower leg*	*N0530*	*Localised osteoarthritis, unspecified, of unspecified site*
N0517	*Localised, primary osteoarthritis of the ankle and foot*	*N0531*	*Localised osteoarthritis, unspecified, of the shoulder region*
N0518	*Localised, primary osteoarthritis of other specified site*	*N0532*	*Localised osteoarthritis, unspecified, of the upper arm*
N0519	*Primary coxarthrosis, bilateral*	*N0533*	*Localised osteoarthritis, unspecified, of the forearm*
N051C	*Primary arthrosis of first carpometacarpal joints, bilateral*	*N0534*	*Localised osteoarthritis, unspecified, of the hand*

CODE	TERM	CODE	TERM
N051D	**Localised, primary osteoarthritis of the wrist**	*N0535*	*Localised osteoarthritis, unspecified, of the pelvic region and thigh*
N053z	*Localised osteoarthritis, unspecified, NOS*	**N05z9**	**Osteoarthritis NOS, of shoulder**
N054	*Oligoarticular osteoarthritis, unspecified*	*N05zA*	*Osteoarthritis NOS, of sternoclavicular joint*
N0540	*Oligoarticular osteoarthritis, unspecified, of unspecified sites*	**N05zE**	**Osteoarthritis NOS, of wrist**
N0541	*Oligoarticular osteoarthritis, unspecified, of the shoulder region*	*N05zF*	*Osteoarthritis NOS, of metacarpophalangeal joint*
N0542	*Oligoarticular osteoarthritis, unspecified, of upper arm*	**N05zG**	**Osteoarthritis NOS, of proximal interphalangeal joint of finger**
N0543	*Oligoarticular osteoarthritis, unspecified, of forearm*	**N05zH**	**Osteoarthritis NOS, of distal interphalangeal joint of finger**
N0544	*Oligoarticular osteoarthritis, unspecified, of hand*	**N05zJ**	**Osteoarthritis NOS, of hip**
N0545	*Oligoarticular osteoarthritis, unspecified, of the pelvic region and thigh*	*N05zK*	*Osteoarthritis NOS, of sacro-iliac joint*
N0546	*Oligoarticular osteoarthritis, unspecified, of lower leg*	**N05zL**	**Osteoarthritis NOS, of knee**
N0547	*Oligoarticular osteoarthritis, unspecified, of ankle and foot*	*N05zM*	*Osteoarthritis NOS, of tibio-fibular joint*
N0548	*Oligoarticular osteoarthritis, unspecified, of other specified sites*	**N05zN**	**Osteoarthritis NOS, of ankle**
N0549	*Oligoarticular osteoarthritis, unspecified, of multiple sites*	**N05zP**	**Osteoarthritis NOS, of subtalar joint**
N054z	*Osteoarthritis of more than one site, unspecified, NOS*	*N05zQ*	*Osteoarthritis NOS, of talonavicular joint*
N05z	**Osteoarthritis NOS**	*N05zR*	*Osteoarthritis NOS, of other tarsal joint*
N05z0	**Osteoarthritis NOS, of unspecified site**	**N05zS**	**Osteoarthritis NOS, of 1st metatarsophalangeal joint**
N05z1	**Osteoarthritis NOS, of shoulder region**	*N05zT*	*Osteoarthritis NOS, of lesser metatarsophalangeal joint*
N05z2	**Osteoarthritis NOS, of the upper arm**	*N05zU*	*Osteoarthritis NOS, of interphalangeal joint of toe*
N05z3	**Osteoarthritis NOS, of the forearm**	**N05zz**	**Osteoarthritis NOS**
N05z4	**Osteoarthritis NOS, of the hand**	**N05zB**	**Osteoarthritis NOS, of acromioclavicular joint**
N05z5	**Osteoarthritis NOS, pelvic region/thigh**	**N05zC**	**Osteoarthritis NOS, of elbow**
N05z6	**Osteoarthritis NOS, of the lower leg**	*N05zD*	*Osteoarthritis NOS, of distal radio-ulnar joint*
N05z7	**Osteoarthritis NOS, of ankle and foot**	**N06z3**	**Arthropathy NOS-forearm**
N05z8	**Osteoarthritis NOS, other specified site**	**N06z4**	**Arthropathy NOS of the hand**
N0536	*Localised osteoarthritis, unspecified, of the lower leg*	**N06z5**	**Hip arthritis NOS**
N0537	*Localised osteoarthritis, unspecified, of the ankle and foot*	**N06z6**	**Knee arthritis NOS**

CODE	TERM	CODE	TERM
N0538	*Localised osteoarthritis, unspecified, of other specified site*	**N06z7**	**Ankle arthritis NOS**
N0539	*Arthrosis of first carpometacarpal joint, unspecified*	**N094**	**Ache in joint**
N0940	**Arthralgia - site unspecified**	**N094H**	**Arthralgia of PIP joint of finger**
N0943	**Arthralgia - forearm**	**N094K**	**Arthralgia of hip**
N0944	**Arthralgia - hand**	**N094M**	**Arthralgia of knee**
N0945	**Arthralgia - pelvic/thigh**	**N094P**	**Arthralgia of ankle**
N0946	**Arthralgia - lower leg**	**N094T**	**Arthralgia of 1st MTP joint**
N0947	**Ankle joint pain**	**N094W**	**Anterior knee pain**
N094F	**Arthralgia of wrist**	**N2450**	**Finger pain**
N094G	**Arthralgia of MCP joint**	**N2451**	**Foot pain**

Bold text indicates that the code appears in the CiPCA database 2007, *italic* that it does not

Appendix E.6. Template joint pain trigger codes (non-N05.. codes): outcome of final MOSAICS team determination of inclusion

Code	Terms	Patients (*n*)	Consultations (*n*)	Retention status	Reason
1M00	Elbow pain	9	14	Exclude	Joint not of interest; considerable non-OA pathology
	Pain in elbow	6	7		
1M10	Knee pain	231	427	Include	Although this includes lots of non-OA pathology, it has usually been included in the RI's OA work
1M11	Foot pain	1	1	Include	
1M13	Ankle pain	1	2	Include	
N065	Polyarthropathy NEC	55	94	Exclude	Much non-OA disease
	Unspecified polyarthropathy	4	6		
N0659	Unsp.polyarthr.-multiple site	3	4	Exclude	Small numbers, infrequent definite OA
N065A	Generalised arthritis	14	22	Exclude	Small numbers, OA pickup low – much possible inflammatory disease or spasticity
N065z	Polyarthritis	12	23	Exclude	Much possible inflammatory disease
N066	Unspecified monoarthritis	1	1	Exclude	Much possible inflammatory disease
N0664	Unsp.monoarthr.-hand	1	3	Exclude	Infrequently used
N06z	Arthritis	38	53	Exclude	Too nonspecific; some child codes retained
	Arthropathy NOS	7	10		
N06z0	Arthropathy NOS-site unspecif.	1	2	Exclude	Infrequently used
N06z1	Arthropathy NOS-shoulder	1	1	Exclude	Joint not of interest; much non-OA pathology
	Shoulder arthritis NOS	7	11		
N06z2	Elbow arthritis NOS	1	1	Exclude	Joint not of interest; much non-OA pathology
N06z3	Arthropathy NOS-forearm	2	4	Include	
	Wrist arthritis NOS	3	3		
N06z4	Arthropathy NOS of the hand	1	1	Include	
	Arthropathy NOS-hand	2	2		
	Hand arthritis NOS	5	7		
N06z5	Hip arthritis NOS	33	55	Include	
N06z6	Knee arthritis NOS	68	129	Include	
N06z7	Ankle arthritis NOS	4	4	Include	
	Foot arthritis NOS	1	1		

Code	Terms	Patients (*n*)	Consultations (*n*)	Retention status	Reason
N06z9	Arthropathy NOS of multiple sites	3	3	Exclude	Small numbers, OA pickup low
N06zA	Acute arthritis	2	2	Exclude	Small numbers, OA pickup low
N06zB	Chronic arthritis	1	1	Exclude	Small numbers, OA pickup low
N06zz	Arthropathy NOS	4	5	Exclude	Small numbers, OA pickup low
N094	Ache in joint	8	10	Include	
	Pain in joint - arthralgia	39	63		
N0940	Arthralgia - site unspecified	8	9	Include	
	Arthralgia of unspecified site	1	2		
N0941	Arthralgia - shoulder	64	100		
	Arthralgia of the shoulder region	61	96	Exclude	Joint not of interest; lots of non-OA pathology
	Shoulder joint pain	67	134		
N0942	Arthralgia - upper arm	1	2	Exclude	Joint not of interest; lots of non-OA pathology
	Elbow joint pain	15	22		
N0943	Arthralgia - forearm	1	1	Include	
	Wrist joint pain	87	125		
N0944	Arthralgia - hand	29	40		
	Arthralgia of the hand	26	41	Include	
	Hand joint pain	13	15		
N0945	Arthralgia - pelvic/thigh	37	61		
	Coxalgia	1	1	Include	
	Hip joint pain	122	184		
N0946	Arthralgia - lower leg	3	4		
	Arthralgia of the lower leg	7	8	Include	
	Knee joint pain	317	460		
N0947	Ankle joint pain	13	19		
	Ankle/foot joint pain	1	1	Include	
	Arthralgia - ankle/foot	31	41		
	Arthralgia of the ankle and foot	6	10		
N0948	Arthralgia - other specified	1	1	Exclude	Small numbers, OA pickup low
N0949	Arthralgia of multiple joints	324	621	Exclude	Lots non-OA pathology
N094A	Arthralgia of shoulder	4	4	Exclude	Joint not of interest; lots of non-OA pathology

Code	Terms	Patients (*n*)	Consultations (*n*)	Retention status	Reason
N094B	Arthralgia - sternoclav joint	3	3	Exclude	Joint not of interest; lots of non-OA pathology
N094C	Arthralgia - acromioclav joint	9	11	Exclude	Joint not of interest; lots of non-OA pathology
N094D	Arthralgia of elbow	23	31	Exclude	Joint not of interest; lots of non-OA pathology
N094F	Arthralgia of wrist	70	111	Include	
N094G	Arthralgia of MCP joint	1	1	Include	
N094H	Arthralgia of PIP joint of finger	1	2	Include	
N094K	Arthralgia of hip	76	118	Include	
	Hip pain	594	955		
N094L	Arthralgia of sacro-iliac joint	5	6	Exclude	Joint not of interest; lots of non-OA pathology
N094M	Arthralgia of knee	636	970	Include	
N094P	Arthralgia of ankle	4	9	Include	
N094T	Arthralgia of 1st MTP joint	5	6	Include	
N094W	Anterior knee pain	209	316	Include	
N094z	Arthralgia NOS	69	108	Exclude	Poor OA specificity
N096	Joint crepitus	2	3	Exclude	Poor OA specificity
	Musculoskeletal pain - joints	46	108		
N0960	Other joint sympt.-site unspec	1	3	Exclude	Poor OA specificity
N0961	Other joint sympt.-shoulder	1	1	Exclude	Joint not of interest; lots of non-OA pathology
N0964	Other joint symptoms of the hand	1	1	Exclude	Only one patient – with a trigger thumb
N0965	Other joint sympt.-pelv./thigh	1	1	Exclude	Poor OA specificity
N0966	Knee gives way	2	4	Exclude	Poor OA specificity
N096B	Other symptoms - sternoclav jt	1	1	Exclude	Joint not of interest; lots of non-OA pathology
N096D	Other symptoms - elbow	2	2	Exclude	Joint not of interest; lots of non-OA pathology
N096F	Other symptoms - wrist	3	3	Exclude	Poor OA specificity

Code	Terms	Patients (*n*)	Consultations (*n*)	Retention status	Reason
N245	Ankle pain	216	333		
	Arm pain	114	170		
	Foot pain	190	280		
	Hand pain	74	108		
	Heel pain	110	161		
	Leg pain	243	357		
	Pain in buttock	32	55	Exclude	Poor OA specificity
	Pain In Left Leg	41	55		
	Pain in limb	15	23		
	Pain in limb - multiple	1	3		
	Pain In Right Arm	29	49		
	Pain In Right Leg	37	49		
	Shoulder pain	353	624		
	Thigh pain	38	55		
N2450	Finger pain	50	71		
	Hand pain	111	161	Include	
	Thumb pain	54	71		
N2451	Foot pain	417	655	Include	
	Toe pain	180	259		

Arthralgia trigger codes: retention decision resulting from MOSAICS team discussions after GP assessment and CiPCA usage filter

Appendix E.7. MOSAICS Template: GP Research Facilitators Meeting 6.4.2011

Background

The MOSAICS study aims to investigate the effects of model consultation training and the provision of a nurse resource on aspects of primary care of osteoarthritis. In order to understand how care changes, we need to be able to compare the intervention practices with controls. Previous evidence suggests that either or both of OA care or its recording were not optimal. To disentangle OA care provision from its recording, we have designed a computer template to use in consultations to capture some of the information that may not routinely be recorded. This template will be implemented in both intervention and control arms of the trial.

Intended template use

The primary aim of the template is not to alter behaviour, rather to facilitate aspects of care which the clinician wishes to record.

Where a clinical feels the template is not applicable, the template may be 'escaped' from - also in situations where there are other reasons not to complete it.

Gaps may be left in the template, though this is not something specifically encouraged.

The template may be completed with the patient in the room, or potentially once the patient has left. The timing of the data entry may affect the responses but this is not a concern.

The template itself

Where a clinician wishes to use the template, once it has been triggered it is a question of entering the code letter for the selected response to each of the domains, with the exception of the numeric value for weight, and the automatic calculation of BMI (assuming a previous height value exists in the system).

The responses are generally quite self-explanatory:

PAIN A Pain|None B Pain|Mild C Pain|Moderate D Pain|Severe	A 'global assessment' which may be the patient's own description or a clinician's estimate based on norm references.
FUNCTION A Fn|Not limited B Fn|Mild limitation C Fn|Moderate limitation D Fn|Severe limitation	Again, a 'global assessment' which may be the patient's own description or a clinician's estimate based on norm references.
WEIGHT	[Numeric value requested]
BMI	[Automatically calculated assuming a previous height record exists]
PARACETAMOL A Para|tried full dose B Para|advised full dose C Para|decline full dose D Para|not appropriate E Para|unknown	This provides an opportunity to record a previous trial of paracetamol – especially since this can be an over-the-counter use which would not otherwise appear in the clinical record. Advice to use paracetamol may also be captured for the same reason, supporting patient self-management. A patient may prefer not to use paracetamol – the reason does not have to be recorded – but this fact can be captured here. It may be considered clinically inappropriate to use paracetamol (e.g. for severe pain where a strong analgesic is considered necessary) so this may be identified (again, no reason is needed). It may be unknown (e.g. if the patient has left before this aspect of the template is completed) and so this option is also provided.
TOPICAL NSAIDs A Top|tried full dose B Top|advised full dose C Top|declined full dose D Top|not appropriate E Top|unknown	This provides an opportunity to record a previous trial of topical NSAIDs – since this can be an over-the-counter use which would not otherwise appear in the clinical record. Advice to use topical NSAIDs may also be captured for the same reason, supporting patient self-management. A patient may prefer not to use topical NSAIDs – the reason does not have to be recorded – but this fact can be captured here. It may be considered clinically inappropriate to use topical NSAIDs (e.g. for severe pain where a strong analgesic is considered necessary) so this may be identified (again, no reason is needed). It may be unknown (e.g. if the patient has left before this aspect of the template is completed) and so this option is also provided.
INFORMATION A Info|verbal B Info|verbal + written C Info|not this time D Info|not appropriate	Provision of information is not always well captured in the record, so this provides an opportunity to record the provision of verbal advice, or verbal + written (on the basis that it would be unlikely to simply hand over a leaflet with no other input). Written advice may refer to a variety of system inbuilt leaflets, special pamphlets, the OA guidebook, etc. The "Not this time" response can be used whenever a consultation does not include information giving, leaving open the possibility that it has previously been given or may be in the future. In some cases it may be considered inappropriate (e.g. in patients with advanced dementia) so an active decision to except a patient may be captured here.

WEIGHT ADVICE A Wt|verbal advice B Wt|verbal + written C Wt|not this time D Wt|not appropriate	Provision of weight advice is not always well captured in the record, so this provides an opportunity to record the provision of verbal advice, or verbal + written. Written advice may refer to a variety of system inbuilt leaflets, special pamphlets, etc. The "Not this time" response can be used whenever a consultation does not include advice on weight management, leaving open the possibility that it has previously been given or may be in the future. In some cases it may be considered inappropriate (e.g. in patients who are not overweight, or where they are not able to participate in weight management for a reason, not specified) so an active decision to except a patient may be captured here.
EXERCISE & PHYSICAL ACTIVITY A Ex|verbal advice B Ex|verbal + written C Ex|not necessary D Ex|not this time E Ex|not appropriate	Similar options are offered for this domain. This provides an opportunity to record the provision of verbal advice, or verbal + written. Written advice may refer to a variety of system inbuilt leaflets, special pamphlets, etc. Some patients may already be exercising to a high level – where exercise information provision is considered not to be needed (superfluous) this may be recorded. The "Not this time" response can be used whenever a consultation does not include advice on exercise/physical activity, leaving open the possibility that it has previously been given or may be in the future. In some cases it may be considered inappropriate (e.g. in patients who are not able to participate for some – unspecified - reason) so an active decision to except a patient may be captured here.
PHYSIOTHERAPY REFERRAL A PT|offered referral B PT|not necessary C PT|not this time D PT|not appropriate	Whenever a physiotherapy referral is offered, or a patient is advised to contact an open access/triage service such as Physio Direct, this is intended to result in the 'Offered referral' response. Some patients may already be exercising to a high level – where physiotherapy referral is considered not to be needed (superfluous) this may be recorded. The "Not this time" response can be used whenever a consultation does not include an offer of a physio referral, leaving open the possibility that it has previously been given or may be in the future. In some cases it may be considered inappropriate (e.g. in patients who are not able to participate for some – unspecified - reason) so an active decision to except a patient may be captured here.

Screenshots

[See Appendix E.3]

Practicalities

The template introduction will be preceded by some training (by John Edwards, at the practice, with a DVD backup for future reference).

The template may be used as soon as it is switched on in a practice.

There will be the opportunity for a review at 3 months with an exchange of ideas between the practice and the study team.

Questions?

John Edwards 6.4.2011

Appendix E.8. Template training DVD script

"Thank you for agreeing to participate in the Keele MOSAICS Study. We aim to describe aspects of osteoarthritis management in general practice, using anonymised data. The best way for us to do this is through the analysis of computerised coded information. We hope that you will be willing to use a computer template, within EMIS, to record information obtained when consulting with patients regarding osteoarthritis.

"I would like to demonstrate the way in which the template can be used. I hope to show that the template is brief, intuitive, and integrates well into a consultation. The template will be triggered by a range of problems and codes, which are considered to reflect the possibility of a working diagnosis of osteoarthritis.

"I would like you to imagine a scenario in which a patient presents with foot pain. In this example, I propose that the diagnosis in this case is biomechanical pain rather than osteoarthritis. Thus, the code 1M11. Foot pain can be entered, which triggers the template as it is a candidate code for a working diagnosis of OA. However, in this scenario, it is considered inappropriate to use the template and so by pressing the 'escape' key, and selecting 'Yes' to the exit without filing prompt, the usual method for recording information can be used.

"In a second scenario, a patient consults with knee pain. This is considered to be due to osteoarthritis. We enter the code 1M10…. Knee pain, and the template fires. This time, the template is completed as it is considered relevant. The same process would be followed if an osteoarthritis disease code (N05..) is entered. Information can be recorded against the system prompts for each of the domains on the single screen shown. For the example, I will complete the template using the topmost option in each domain.

"Thus for pain, the severity can be recorded against the broad groups offered – ranging from 'No pain' through 'Mild' and 'Moderate' to 'Severe pain'.

"For recording information on the degree of functional impairment, a summary assessment can be coded with options ranging from 'Not limited' through 'Mild' and 'Moderate' to 'Severe limitation'.

“For weight, the numeric value entered will allow the system to calculate BMI automatically if a patient’s height has previously been recorded.

“A previous trial of paracetamol at a dose appropriate to that patient can be recorded, as can advice to the patient to try paracetamol up to an appropriate maximal dose. If a patient’s preference is not to use paracetamol, this may be coded as “Declined”. If such a trial is not considered to be clinically appropriate, the response “Not appropriate” may be used. “Unknown” is also an option, if required.

“The same options are provided for topical anti-inflammatory agents. A previous trial of topical anti-inflammatories at a dose appropriate to that patient can be recorded, as can advice to the patient to try topical anti-inflammatories up to an appropriate dose. If a patient’s preference is not to use topical anti-inflammatories, this may be coded as “Declined”. If such a trial is not considered to be clinically appropriate, the response “Not appropriate” may be used. “Unknown” is also an option, if required.

“Provision of information about osteoarthritis may be recorded as ‘Verbal advice’, or ‘Verbal and written’. If information is not given in a consultation, that can also be recorded as “Not this time”. In some circumstances, information provision may be considered not to be appropriate, and this might be noted too.

“Similarly, advice to the patient about weight management may be recorded as ‘Verbal’ or ‘Verbal and written’. If such advice is not given in a consultation, this may be recorded as “Not this time.” “It may be that provision of weight advice is considered not to be appropriate and this may also be recorded.

“The same options of ‘Verbal’ or ‘Verbal and written’ are offered for recording advice about exercise and physical activity, with the addition of a “Not necessary” category. This may be used if a patient is already undertaking exercise and physical activity deemed to be of an appropriate type, intensity, and frequency. The other options of ‘Not this time’ and ‘Not appropriate’ remain available.

“Likewise, physiotherapy referral may be considered and recorded. In this case, it is the offer of a referral or advice to contact an open-access or triage service that should trigger the “Offered referral” prompt, not only cases in which the referral was accepted. In some cases, Physio referral may not be considered necessary, such as in patients who are already well aware of the type of exercise to do: this may be recorded as “Not necessary”. If a referral is not offered, or has previously been recommended and not reiterated or it is not a current treatment strand, this should be recorded as “Not this time”. “Not appropriate” can be used in circumstances where such a referral is not a clinically appropriate course of action, such as in someone unable to participate in a physiotherapy treatment.

“At completion of the template, the information can be filed [<Return>] and the record then used as normal to record any other information necessary in the usual data entry manner.

“This has been a quick demonstration of the template. I hope that it has demonstrated a practicable way of using it and that you will feel able to complete the template as appropriate in clinical practice. You may have additional questions. These are can be put to the Keele Informatics team – contact Tracy Whitehurst on 01782 734712, t.whitehurst@cphc.keele.ac.uk”

Appendix E.9. Mapping of Read codes to sites of disease

Read code	Term	Site allocation	Defined OA or joint pain	Reclassified OA 'not OA' or joint pain
1M10	Knee pain	Knee	Joint pain	OA
1M11	Foot pain	Ankle/foot	Joint pain	Not OA
1M13	Ankle pain	Ankle/foot	Joint pain	Not OA
EGTON279	Painful Right Knee	Knee	Joint pain	OA
N05	Osteoarthritis+allied disord.	Unspecified	OA	OA
N05-1	Osteoarthritis	Unspecified	OA	OA
N050	Generalised osteoarthritis - OA	Generalised	OA	OA
N0501	Generalised OA-hand	Wrist/hand	OA	OA
N0501-1	Heberdens' nodes	Wrist/hand	OA	OA
N0502	Generalised OA-multiple sites	Generalised	OA	OA
N0502-99	Osteoarthritis -Multiple Joint	Multiple	OA	OA
N0514	Local.primary OA-hand	Wrist/hand	OA	OA
N0517	Local.primary OA-ankle/foot	Ankle/foot	OA	OA
N0519	Primary coxarthrosis bilateral	Hip	OA	OA
N051E	Local prim osteoarth toe	Ankle/foot	OA	OA
N0536-1	Patellofemoral osteoarthritis	Knee	OA	OA
N0537	Local.OA unsp.-ankle/foot	Ankle/foot	OA	OA
N05z	Osteoarthritis NOS	Unspecified	OA	OA
N05z1	Osteoarthritis NOS-shoulder	Shoulder	OA	OA
N05z1-99	Osteoarthritis -Shoulder Joint	Shoulder	OA	OA
N05z2-1	Elbow osteoarthritis NOS	Elbow	OA	OA
N05z3-1	Wrist osteoarthritis NOS	Wrist/hand	OA	OA
N05z3-99	Osteoarthritis - Wrist Joint	Wrist/hand	OA	OA
N05z4	Osteoarthritis NOS of the hand	Wrist/hand	OA	OA
N05z4-1	Finger osteoarthritis NOS	Wrist/hand	OA	OA
N05z4-2	Thumb osteoarthritis NOS	Wrist/hand	OA	OA
N05z4-99	Osteoarthritis - Hand Joint	Wrist/hand	OA	OA
N05z5-1	Hip osteoarthritis NOS	Hip	OA	OA
N05z5-99	Osteoarthritis - Hip Joint	Hip	OA	OA
N05z6	Osteoarthritis NOS-lower leg	Knee	OA	OA
N05z6-1	Knee osteoarthritis NOS	Knee	OA	OA
N05z6-99	Osteoarthritis - Knee Joint	Knee	OA	OA
N05z7	Osteoarthritis NOS-ankle/foot	Ankle/foot	OA	OA
N05z7-1	Ankle osteoarthritis NOS	Ankle/foot	OA	OA
N05z7-2	Foot osteoarthritis NOS	Ankle/foot	OA	OA
N05z7-3	Toe osteoarthritis NOS	Ankle/foot	OA	OA
N05z7-99	Osteoarthritis - Ankle/Foot	Ankle/foot	OA	OA
N05z8	Osteoarthritis NOS-other spec	Unspecified	OA	OA
N05z8-99	Osteoarthritis - Other Joint	Unspecified	OA	OA
N05z9	Osteoarthritis NOS of shoulder	Shoulder	OA	OA
N05zB	OA NOS-acromioclavicular joint	Shoulder	OA	OA
N05zE	Osteoarthritis NOS of wrist	Wrist/hand	OA	OA
N05zF	Osteoarthritis NOS of MCP joint	Wrist/hand	OA	OA
N05zJ	Osteoarthritis NOS of hip	Hip	OA	OA
N05zL	Osteoarthritis NOS of knee	Knee	OA	OA
N05zS	OA NOS-1st MTP joint	Ankle/foot	OA	OA
N05zz	Osteoarthritis NOS	Unspecified	OA	OA
N06z3-1	Wrist arthritis NOS	Wrist/hand	Joint pain	Not OA
N06z4	Arthropathy NOS of the hand	Wrist/hand	Joint pain	Not OA
N06z4-1	Hand arthritis NOS	Wrist/hand	Joint pain	Not OA
N06z5-1	Hip arthritis NOS	Hip	Joint pain	OA
N06z6-1	Knee arthritis NOS	Knee	Joint pain	OA
N06z7-1	Ankle arthritis NOS	Ankle/foot	Joint pain	Not OA
N06z7-2	Foot arthritis NOS	Ankle/foot	Joint pain	Not OA

Read code	Term	Site allocation	Defined OA or joint pain	Reclassified OA 'not OA' or joint pain
N094	Pain in joint - arthralgia	Unspecified	Joint pain	OA
N0940	Arthralgia of unspecified site	Unspecified	Joint pain	OA
N0943	Arthralgia - forearm	Wrist/hand	Joint pain	Not OA
N0943-1	Wrist joint pain	Wrist/hand	Joint pain	Not OA
N0943-99	Wrist Joint Pain	Wrist/hand	Joint pain	Not OA
N0944	Arthralgia of the hand	Wrist/hand	Joint pain	Not OA
N0944-1	Hand joint pain	Wrist/hand	Joint pain	Not OA
N0944-99	Hand Joint Pain	Wrist/hand	Joint pain	Not OA
N0945	Arthralgia - pelvic/thigh	Hip	Joint pain	OA
N0945-2	Hip joint pain	Hip	Joint pain	OA
N0945-99	Hip Joint Pain	Hip	Joint pain	OA
N0946	Arthralgia - lower leg	Knee	Joint pain	OA
N0946-1	Knee joint pain	Knee	Joint pain	OA
N0946-99	Knee Joint Pain	Knee	Joint pain	OA
N0947	Arthralgia - ankle/foot	Ankle/foot	Joint pain	Not OA
N0947-1	Ankle joint pain	Ankle/foot	Joint pain	Not OA
N0947-99	Ankle/Foot Joint Pain	Ankle/foot	Joint pain	Not OA
N094F	Arthralgia of wrist	Wrist/hand	Joint pain	Not OA
N094F-1	Wrist pain	Wrist/hand	Joint pain	Not OA
N094G	Arthralgia of MCP joint	Wrist/hand	Joint pain	Not OA
N094H	Arthralgia of PIP joint of finger	Wrist/hand	Joint pain	Not OA
N094K	Arthralgia of hip	Hip	Joint pain	OA
N094K-2	Hip pain	Hip	Joint pain	OA
N094M	Arthralgia of knee	Knee	Joint pain	OA
N094P	Arthralgia of ankle	Ankle/foot	Joint pain	Not OA
N094T	Arthralgia of 1st MTP joint	Ankle/foot	Joint pain	Not OA
N094W	Anterior knee pain	Knee	Joint pain	OA
N2450	Hand pain	Wrist/hand	Joint pain	Not OA
N2450-1	Thumb pain	Wrist/hand	Joint pain	Not OA
N2450-2	Finger pain	Wrist/hand	Joint pain	Not OA
N2451	Foot pain	Ankle/foot	Joint pain	Not OA
N2451-1	Toe pain	Ankle/foot	Joint pain	Not OA

Read codes mapped to site of disease with initial and revised suggestions for use as a definition of clinical OA

Appendix F. Routinely-recorded quality of care appendices (Chapter Five): Associations of routinely-recorded measures with independent variables

Appendix F.1. Assessment processes

Appendix F.1.1. Weight record

Weight record within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.26 (0.97,1.63)	1.00 (0.75,1.34)	1.00 (0.74,1.34)
Sex (*reference*: female)	1.21 (0.95,1.54)	1.26 (0.98,1.62)	1.25 (0.97,1.61)
Age 65-74 (*reference*: 45-64)	1.00 (0.76,1.33)	0.89 (0.66,1.19)	0.89 (0.66,1.20)
Age 75-84	0.74 (0.53,1.04)	0.64 (0.44,0.93)	0.64 (0.44,0.92)
Age 85+	0.32 (0.16,0.64)	0.31 (0.15,0.63)	0.30 (0.15,0.63)
Hip (*reference*: knee)	0.82 (0.59,1.14)	0.88 (0.63,1.23)	0.88 (0.62,1.23)
Ankle/foot	0.72 (0.41,1.25)	0.85 (0.48,1.49)	0.84 (0.47,1.48)
Wrist/hand	0.87 (0.55,1.38)	1.05 (0.65,1.69)	1.04 (0.64,1.68)
Unspecified	0.81 (0.44,1.49)	0.91 (0.48,1.71)	0.93 (0.49,1.75)
Multisite	1.60 (1.12,2.29)	1.22 (0.83,1.79)	1.21 (0.82,1.79)
BMI 25 to <30 (reference: BMI <25)	1.07 (0.72,1.60)	0.93 (0.62,1.39)	0.94 (0.62,1.41)
BMI 30+	1.79 (1.24,2.59)	1.40 (0.96,2.05)	1.42 (0.97,2.09)
BMI unknown	0.47 (0.30,0.73)	0.41 (0.26,0.64)	0.41 (0.26,0.64)
5-9 BNF chapters (*reference*: 0-4)	1.20 (0.90,1.59)	1.05 (0.77,1.42)	1.05 (0.77,1.43)
10+ BNF chapters	1.43 (1.06,1.93)	1.35 (0.96,1.89)	1.34 (0.96,1.88)
Multiple consultations in phase 1 (vs. single)	2.52 (1.97,3.23)	2.38 (1.83,3.10)	2.36 (1.81,3.08)
X-ray, phase 1[a] (vs. none)	1.68 (1.11,2.52)	1.39 (0.89,2.16)	1.33 (0.86,2.07)
Above the clinician median index consultation count (*reference*: at or below the median)	1.21 (0.68,2.15)	1.34 (0.74,2.40)	1.41 (0.79,2.53)

Weight record within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.37 (0.64,2.96)	1.52 (0.73,3.16)	-
Practice 3	0.46 (0.25,0.86)	0.53 (0.29,0.97)	-
Practice 4	1.22 (0.74,2.01)	1.59 (0.98,2.60)	-
Practice 5	1.25 (0.73,2.15)	1.29 (0.77,2.17)	-
Practice 6	2.08 (1.03,4.19)	1.88 (0.96,3.67)	-
Practice 7	0.93 (0.53,1.64)	0.93 (0.53,1.64)	-
Practice 8	1.27 (0.61,2.63)	1.04 (0.51,2.14)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between a weight record within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.1.2. X-ray record

X-ray record within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	0.78 (0.62,0.97)	0.54 (0.41,0.71)	0.54 (0.41,0.71)
Sex (*reference*: female)	0.91 (0.74,1.13)	0.88 (0.70,1.10)	0.87 (0.70,1.10)
Age 65-74 (*reference*: 45-64)	1.07 (0.84,1.36)	1.15 (0.88,1.51)	1.15 (0.88,1.51)
Age 75-84	0.99 (0.75,1.31)	1.07 (0.78,1.47)	1.08 (0.78,1.48)
Age 85+	0.42 (0.26,0.68)	0.48 (0.28,0.82)	0.49 (0.28,0.83)
Hip (*reference*: knee)	1.73 (1.35,2.21)	1.92 (1.46,2.51)	1.92 (1.44,2.55)
Ankle/foot	0.45 (0.26,0.79)	0.54 (0.31,0.95)	0.54 (0.30,0.95)
Wrist/hand	0.51 (0.33,0.79)	0.55 (0.35,0.88)	0.55 (0.34,0.88)
Unspecified	0.29 (0.14,0.61)	0.39 (0.18,0.86)	0.39 (0.18,0.87)
Multisite	1.20 (0.87,1.65)	1.01 (0.70,1.45)	1.03 (0.72,1.48)
BMI 25 to <30 (reference: BMI <25)	1.58 (1.11,2.24)	1.61 (1.10,2.34)	1.60 (1.09,2.34)
BMI 30+	1.39 (0.99,1.94)	1.26 (0.87,1.81)	1.26 (0.87,1.81)
BMI unknown	1.13 (0.81,1.59)	1.09 (0.76,1.57)	1.09 (0.76,1.57)
5-9 BNF chapters (*reference*: 0-4)	1.05 (0.83,1.33)	0.90 (0.69,1.18)	0.91 (0.69,1.18)
10+ BNF chapters	0.80 (0.61,1.05)	0.63 (0.46,0.87)	0.64 (0.46,0.89)
Multiple consultations in phase 1 (vs. single)	4.56 (3.66,5.67)	4.99 (3.95,6.31)	4.96 (3.51,7.00)
X-ray, phase 1[a] (vs. none)	2.41 (1.79,3.26)	1.82 (1.30,2.55)	1.78 (1.25,2.53)
Above the clinician median index consultation count (*reference*: at or below the median)	0.54 (0.25,1.15)	1.06 (0.69,1.65)	1.04 (0.66,1.64)

X-ray record within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.71 (0.38,1.31)	0.90 (0.46,1.78)	-
Practice 3	0.18 (0.11,0.29)	0.12 (0.07,0.21)	-
Practice 4	0.02 (0.00,0.05)	0.01 (0.00,0.05)	-
Practice 5	0.02 (0.01,0.06)	0.01 (0.00,0.05)	-
Practice 6	0.02 (0.00,0.13)	0.01 (0.00,0.08)	-
Practice 7	1.53 (1.05,2.23)	1.42 (0.92,2.19)	-
Practice 8	1.13 (0.67,1.90)	0.83 (0.46,1.50)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between a relevant recorded X-ray within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.2. Pharmacological management

Appendix F.2.1. Paracetamol

Paracetamol prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	2.21 (1.79,2.73)	1.58 (1.23,2.03)	1.57 (1.22,2.02)
Sex (*reference*: female)	0.70 (0.56,0.87)	0.76 (0.60,0.95)	0.76 (0.60,0.96)
Age 65-74 (*reference*: 45-64)	2.42 (1.82,3.21)	1.93 (1.44,2.60)	1.94 (1.44,2.60)
Age 75-84	4.60 (3.48,6.09)	3.29 (2.43,4.45)	3.27 (2.42,4.44)
Age 85+	5.43 (3.80,7.78)	3.75 (2.53,5.55)	3.73 (2.52,5.52)
Hip (*reference*: knee)	1.48 (1.14,1.90)	1.37 (1.04,1.79)	1.37 (1.05,1.80)
Ankle/foot	0.43 (0.23,0.78)	0.51 (0.27,0.96)	0.52 (0.28,0.98)
Wrist/hand	0.49 (0.30,0.82)	0.59 (0.35,1.00)	0.60 (0.35,1.01)
Unspecified	1.40 (0.87,2.24)	1.22 (0.74,2.03)	1.20 (0.72,1.99)
Multisite	2.12 (1.57,2.88)	1.19 (0.84,1.66)	1.21 (0.86,1.69)
BMI 25 to <30 (reference: BMI <25)	1.02 (0.73,1.42)	1.07 (0.75,1.52)	1.05 (0.73,1.50)
BMI 30+	1.07 (0.78,1.48)	1.08 (0.76,1.53)	1.07 (0.75,1.52)
BMI unknown	0.79 (0.57,1.10)	0.93 (0.66,1.32)	0.92 (0.65,1.31)
5-9 BNF chapters (*reference*: 0-4)	2.45 (1.90,3.17)	1.62 (1.23,2.13)	1.62 (1.23,2.14)
10+ BNF chapters	3.79 (2.92,4.93)	2.04 (1.53,2.74)	2.06 (1.53,2.76)
Multiple consultations in phase 1 (vs. single)	2.14 (1.74,2.63)	1.83 (1.46,2.30)	1.84 (1.46,2.31)
X-ray, phase 1[a] (vs. none)	1.36 (0.95,1.96)	0.95 (0.64,1.41)	0.95 (0.64,1.42)
Above the clinician median index consultation count (*reference*: at or below the median)	1.22 (0.78,1.93)	1.01 (0.62,1.65)	1.04 (0.64,1.69)

Paracetamol prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.62 (0.35,1.12)	0.76 (0.41,1.43)	-
Practice 3	0.94 (0.66,1.32)	1.20 (0.82,1.75)	-
Practice 4	0.51 (0.34,0.75)	0.53 (0.35,0.81)	-
Practice 5	0.66 (0.45,0.98)	0.90 (0.59,1.37)	-
Practice 6	0.90 (0.53,1.53)	1.10 (0.62,1.97)	-
Practice 7	0.63 (0.42,0.95)	0.80 (0.51,1.26)	-
Practice 8	0.62 (0.35,1.11)	0.55 (0.29,1.04)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between paracetamol prescription within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.2.2. Topical NSAIDs

Topical NSAIDs prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.38 (1.12,1.71)	0.97 (0.75,1.24)	0.98 (0.77,1.26)
Sex (*reference*: female)	0.67 (0.54,0.82)	0.67 (0.53,0.83)	0.66 (0.53,0.83)
Age 65-74 (*reference*: 45-64)	2.47 (1.93,3.17)	2.45 (1.88,3.19)	2.45 (1.88,3.20)
Age 75-84	2.54 (1.93,3.33)	2.32 (1.72,3.13)	2.33 (1.72,3.14)
Age 85+	3.02 (2.08,4.38)	2.71 (1.80,4.07)	2.68 (1.78,4.04)
Hip (*reference*: knee)	0.40 (0.30,0.55)	0.34 (0.25,0.47)	0.34 (0.24,0.46)
Ankle/foot	0.52 (0.32,0.83)	0.53 (0.32,0.88)	0.53 (0.32,0.88)
Wrist/hand	1.11 (0.79,1.57)	1.27 (0.89,1.82)	1.28 (0.89,1.83)
Unspecified	0.97 (0.62,1.54)	0.92 (0.56,1.50)	0.91 (0.56,1.49)
Multisite	1.39 (1.02,1.90)	1.03 (0.73,1.46)	1.02 (0.73,1.44)
BMI 25 to <30 (reference: BMI <25)	1.15 (0.83,1.62)	1.16 (0.82,1.65)	1.16 (0.82,1.66)
BMI 30+	1.09 (0.78,1.51)	1.04 (0.73,1.47)	1.04 (0.73,1.47)
BMI unknown	1.06 (0.77,1.46)	1.23 (0.87,1.72)	1.23 (0.87,1.73)
5-9 BNF chapters (*reference*: 0-4)	1.78 (1.40,2.26)	1.40 (1.08,1.82)	1.41 (1.09,1.83)
10+ BNF chapters	2.44 (1.90,3.12)	1.82 (1.37,2.41)	1.83 (1.38,2.43)
Multiple consultations in phase 1 (vs. single)	1.71 (1.40,2.09)	1.59 (1.27,1.98)	1.59 (1.27,1.99)
X-ray, phase 1[a] (vs. none)	0.96 (0.65,1.43)	1.00 (0.66,1.53)	0.99 (0.65,1.52)
Above the clinician median index consultation count (*reference*: at or below the median)	0.95 (0.59,1.54)	0.75 (0.44,1.25)	0.77 (0.46,1.30)

Topical NSAIDs prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	3.47 (1.52,7.94)	4.61 (1.87,11.4)	-
Practice 3	1.09 (0.64,1.86)	1.20 (0.67,2.14)	-
Practice 4	1.14 (0.65,1.99)	1.21 (0.66,2.24)	-
Practice 5	2.17 (1.23,3.82)	2.59 (1.39,4.83)	-
Practice 6	0.71 (0.28,1.78)	0.74 (0.28,1.99)	-
Practice 7	0.70 (0.38,1.28)	0.68 (0.35,1.32)	-
Practice 8	2.47 (1.25,4.88)	2.56 (1.19,5.51)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between topical NSAID prescription within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.2.3. Paracetamol or topical NSAIDs

Paracetamol or topical NSAID prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.85 (1.56,2.20)	1.27 (1.03,1.55)	1.29 (1.05,1.58)
Sex (*reference*: female)	0.63 (0.53,0.75)	0.65 (0.54,0.78)	0.65 (0.54,0.78)
Age 65-74 (*reference*: 45-64)	2.57 (2.09,3.17)	2.29 (1.84,2.85)	2.28 (1.83,2.84)
Age 75-84	3.78 (3.03,4.72)	3.04 (2.39,3.88)	3.03 (2.38,3.86)
Age 85+	4.76 (3.51,6.47)	3.75 (2.69,5.23)	3.69 (2.65,5.15)
Hip (*reference*: knee)	0.81 (0.66,1.01)	0.69 (0.55,0.87)	0.68 (0.54,0.86)
Ankle/foot	0.47 (0.32,0.71)	0.51 (0.34,0.79)	0.52 (0.34,0.80)
Wrist/hand	0.85 (0.63,1.16)	0.99 (0.71,1.37)	0.99 (0.72,1.37)
Unspecified	1.35 (0.93,1.95)	1.22 (0.81,1.83)	1.20 (0.80,1.80)
Multisite	1.77 (1.37,2.29)	1.14 (0.85,1.52)	1.13 (0.84,1.51)
BMI 25 to <30 (reference: BMI <25)	1.06 (0.81,1.39)	1.10 (0.83,1.47)	1.10 (0.83,1.47)
BMI 30+	1.10 (0.85,1.42)	1.09 (0.82,1.44)	1.09 (0.82,1.44)
BMI unknown	0.87 (0.67,1.13)	1.04 (0.79,1.38)	1.04 (0.79,1.37)
5-9 BNF chapters (*reference*: 0-4)	2.18 (1.79,2.65)	1.53 (1.24,1.89)	1.54 (1.24,1.90)
10+ BNF chapters	3.45 (2.81,4.25)	2.11 (1.67,2.67)	2.13 (1.68,2.68)
Multiple consultations in phase 1 (vs. single)	1.88 (1.60,2.22)	1.69 (1.41,2.03)	1.69 (1.41,2.03)
X-ray, phase 1[a] (vs. none)	1.16 (0.85,1.57)	1.00 (0.71,1.40)	0.99 (0.71,1.39)
Above the clinician median index consultation count (*reference*: at or below the median)	1.06 (0.73,1.53)	0.85 (0.56,1.28)	0.88 (0.58,1.32)

Paracetamol or topical NSAID prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.75 (0.95,3.23)	2.27 (1.15,4.51)	-
Practice 3	0.92 (0.63,1.36)	1.10 (0.72,1.70)	-
Practice 4	0.76 (0.51,1.13)	0.78 (0.49,1.23)	-
Practice 5	1.32 (0.87,2.00)	1.76 (1.09,2.82)	-
Practice 6	0.87 (0.47,1.61)	0.98 (0.49,1.94)	-
Practice 7	0.65 (0.42,1.00)	0.71 (0.44,1.15)	-
Practice 8	1.47 (0.86,2.50)	1.48 (0.81,2.72)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between paracetamol or topical NSAID prescription within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.2.4. Opioids

Opioids prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	2.06 (1.75,2.43)	1.44 (1.19,1.75)	1.46 (1.20,1.77)
Sex (*reference*: female)	0.82 (0.70,0.95)	0.88 (0.74,1.04)	0.88 (0.74,1.05)
Age 65-74 (*reference*: 45-64)	1.69 (1.40,2.03)	1.38 (1.13,1.70)	1.39 (1.13,1.70)
Age 75-84	1.90 (1.54,2.33)	1.33 (1.05,1.69)	1.32 (1.04,1.68)
Age 85+	2.43 (1.81,3.26)	1.72 (1.24,2.39)	1.71 (1.23,2.38)
Hip (*reference*: knee)	1.33 (1.10,1.61)	1.36 (1.11,1.68)	1.36 (1.11,1.68)
Ankle/foot	0.36 (0.24,0.55)	0.43 (0.28,0.65)	0.42 (0.27,0.64)
Wrist/hand	0.46 (0.33,0.63)	0.56 (0.39,0.79)	0.56 (0.40,0.79)
Unspecified	1.26 (0.88,1.81)	1.20 (0.81,1.80)	1.20 (0.80,1.79)
Multisite	2.17 (1.69,2.78)	1.13 (0.86,1.49)	1.12 (0.85,1.48)
BMI 25 to <30 (reference: BMI <25)	1.03 (0.79,1.33)	1.00 (0.76,1.33)	1.00 (0.75,1.32)
BMI 30+	1.64 (1.28,2.09)	1.44 (1.10,1.89)	1.44 (1.10,1.89)
BMI unknown	0.93 (0.73,1.19)	1.02 (0.78,1.34)	1.01 (0.77,1.33)
5-9 BNF chapters (*reference*: 0-4)	2.09 (1.74,2.50)	1.62 (1.32,1.98)	1.63 (1.34,2.00)
10+ BNF chapters	3.67 (3.01,4.48)	2.66 (2.13,3.34)	2.72 (2.17,3.41)
Multiple consultations in phase 1 (vs. single)	3.40 (2.90,3.98)	3.05 (2.57,3.63)	3.07 (2.58,3.65)
X-ray, phase 1[a] (vs. none)	1.63 (1.23,2.15)	1.04 (0.76,1.42)	1.04 (0.76,1.43)
Above the clinician median index consultation count (*reference*: at or below the median)	0.68 (0.49,0.94)	0.71 (0.50,1.01)	0.68 (0.48,0.98)

Opioids prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.89 (1.23,2.91)	2.42 (1.45,4.06)	-
Practice 3	0.83 (0.62,1.11)	0.97 (0.69,1.37)	-
Practice 4	0.59 (0.43,0.79)	0.66 (0.46,0.94)	-
Practice 5	0.71 (0.52,0.97)	0.95 (0.65,1.38)	-
Practice 6	1.32 (0.86,2.03)	1.45 (0.87,2.44)	-
Practice 7	1.52 (1.13,2.05)	1.81 (1.26,2.60)	-
Practice 8	1.37 (0.90,2.07)	1.13 (0.69,1.85)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between opioid prescription within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.2.5. Oral NSAIDs

Oral NSAIDs prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.16 (0.95,1.43)	1.03 (0.81,1.31)	1.03 (0.81,1.31)
Sex (*reference*: female)	1.13 (0.93,1.37)	1.10 (0.90,1.35)	1.10 (0.90,1.35)
Age 65-74 (*reference*: 45-64)	0.62 (0.49,0.78)	0.62 (0.49,0.79)	0.62 (0.48,0.79)
Age 75-84	0.45 (0.34,0.61)	0.46 (0.34,0.62)	0.46 (0.33,0.62)
Age 85+	0.22 (0.13,0.39)	0.22 (0.12,0.39)	0.22 (0.12,0.39)
Hip (*reference*: knee)	0.95 (0.74,1.21)	1.06 (0.82,1.38)	1.06 (0.82,1.38)
Ankle/foot	0.87 (0.57,1.35)	1.01 (0.65,1.58)	1.01 (0.64,1.58)
Wrist/hand	0.92 (0.64,1.33)	1.10 (0.75,1.60)	1.09 (0.74,1.60)
Unspecified	1.14 (0.72,1.78)	1.36 (0.85,2.19)	1.38 (0.85,2.22)
Multisite	1.55 (1.15,2.08)	1.24 (0.89,1.72)	1.24 (0.89,1.72)
BMI 25 to <30 (reference: BMI <25)	1.32 (0.94,1.85)	1.26 (0.89,1.79)	1.26 (0.89,1.80)
BMI 30+	1.31 (0.94,1.81)	1.14 (0.81,1.61)	1.14 (0.81,1.61)
BMI unknown	1.51 (1.10,2.07)	1.33 (0.96,1.85)	1.33 (0.96,1.86)
5-9 BNF chapters (*reference*: 0-4)	0.87 (0.70,1.08)	0.99 (0.78,1.26)	1.00 (0.78,1.27)
10+ BNF chapters	0.59 (0.45,0.77)	0.73 (0.54,0.98)	0.73 (0.54,0.99)
Multiple consultations in phase 1 (vs. single)	2.56 (2.11,3.11)	2.61 (2.12,3.22)	2.62 (2.12,3.24)
X-ray, phase 1[a] (vs. none)	1.80 (1.31,2.47)	1.38 (0.97,1.96)	1.38 (0.97,1.97)
Above the clinician median index consultation count (*reference*: at or below the median)	0.95 (0.60,1.50)	0.90 (0.58,1.41)	0.93 (0.59,1.47)

Oral NSAIDs prescription within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.83 (0.39,1.80)	0.90 (0.42,1.92)	-
Practice 3	1.22 (0.78,1.90)	1.19 (0.76,1.85)	-
Practice 4	0.75 (0.46,1.23)	0.87 (0.54,1.42)	-
Practice 5	0.37 (0.21,0.67)	0.35 (0.20,0.63)	-
Practice 6	0.89 (0.43,1.86)	0.91 (0.44,1.89)	-
Practice 7	0.34 (0.19,0.62)	0.29 (0.16,0.52)	-
Practice 8	0.68 (0.34,1.35)	0.59 (0.29,1.17)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between oral NSAID prescription within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.2.6. Oral NSAIDs in the presence of a recorded relative contraindication

Oral NSAIDs prescription within 14 days of a clinical OA consultation with relative contraindication, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
OA diagnosis (*reference:* joint pain)	1.36 (1.00,1.85)	1.18 (0.82,1.69)	1.19 (0.83,1.71)
Sex (*reference*: female)	1.11 (0.83,1.50)	1.14 (0.83,1.55)	1.13 (0.83,1.55)
Age 65-74 (*reference*: 45-64)	0.54 (0.38,0.77)	0.53 (0.37,0.76)	0.52 (0.36,0.76)
Age 75-84	0.41 (0.27,0.61)	0.37 (0.24,0.56)	0.37 (0.24,0.57)
Age 85+	0.33 (0.17,0.64)	0.29 (0.15,0.58)	0.29 (0.15,0.58)
Hip (*reference*: knee)	0.94 (0.64,1.39)	1.03 (0.69,1.53)	1.02 (0.68,1.53)
Ankle/foot	1.11 (0.60,2.05)	1.29 (0.69,2.41)	1.29 (0.68,2.43)
Wrist/hand	1.00 (0.56,1.77)	1.23 (0.68,2.21)	1.22 (0.67,2.21)
Unspecified	0.99 (0.44,2.20)	1.24 (0.54,2.86)	1.21 (0.52,2.80)
Multisite	1.72 (1.13,2.62)	1.37 (0.86,2.17)	1.37 (0.86,2.18)
BMI 25 to <30 (reference: BMI <25)	1.62 (0.96,2.74)	1.49 (0.88,2.55)	1.51 (0.88,2.59)
BMI 30+	1.62 (0.98,2.68)	1.28 (0.76,2.16)	1.27 (0.75,2.16)
BMI unknown	1.39 (0.80,2.42)	1.18 (0.67,2.08)	1.20 (0.68,2.12)
5-9 BNF chapters (*reference*: 0-4)	0.95 (0.67,1.35)	1.09 (0.75,1.59)	1.11 (0.76,1.61)
10+ BNF chapters	0.91 (0.63,1.32)	1.10 (0.73,1.64)	1.10 (0.73,1.66)
Multiple consultations in phase 1 (vs. single)	2.24 (1.66,3.02)	2.00 (1.44,2.76)	2.03 (1.46,2.81)
X-ray, phase 1[a] (vs. none)	2.16 (1.39,3.36)	1.60 (0.99,2.60)	1.60 (0.99,2.61)
Above the clinician median index consultation count (*reference*: at or below the median)	1.16 (0.59,2.26)	1.06 (0.55,2.05)	1.09 (0.56,2.14)

Oral NSAIDs prescription within 14 days of a clinical OA consultation with relative contraindication, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.57 (0.23,1.45)	0.67 (0.26,1.71)	-
Practice 3	1.64 (1.03,2.61)	1.78 (1.09,2.89)	-
Practice 4	0.66 (0.35,1.24)	0.83 (0.43,1.58)	-
Practice 5	0.41 (0.21,0.81)	0.44 (0.23,0.87)	-
Practice 6	1.24 (0.63,2.43)	1.27 (0.63,2.55)	-
Practice 7	0.25 (0.11,0.55)	0.24 (0.10,0.55)	-
Practice 8	0.76 (0.36,1.63)	0.74 (0.34,1.61)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between oral NSAID prescription in the presence of a relative comorbidity contraindication, within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.2.7. PPIs (in those prescribed oral NSAIDs)

PPI prescription within 14 days of a clinical OA consultation, phase 1 (in those prescribed an oral NSAID)	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
OA diagnosis (*reference:* joint pain)	1.34 (0.93,1.93)	1.06 (0.67,1.68)	1.11 (0.71,1.75)
Sex (*reference*: female)	0.72 (0.50,1.04)	0.71 (0.48,1.05)	0.74 (0.50,1.09)
Age 65-74 (*reference*: 45-64)	2.02 (1.32,3.10)	2.11 (1.33,3.35)	2.17 (1.37,3.45)
Age 75-84	2.64 (1.55,4.49)	2.16 (1.19,3.90)	2.08 (1.15,3.77)
Age 85+	12.1 (3.22,45.7)	9.66 (2.35,39.6)	10.9 (2.63,44.8)
Hip (*reference*: knee)	1.22 (0.77,1.93)	1.10 (0.66,1.86)	1.13 (0.67,1.89)
Ankle/foot	1.18 (0.52,2.66)	1.23 (0.51,2.97)	1.32 (0.55,3.15)
Wrist/hand	0.80 (0.38,1.66)	0.87 (0.40,1.92)	0.88 (0.40,1.93)
Unspecified	1.24 (0.55,2.82)	0.97 (0.39,2.41)	0.93 (0.38,2.28)
Multisite	1.03 (0.61,1.75)	0.65 (0.35,1.21)	0.61 (0.33,1.12)
BMI 25 to <30 (reference: BMI <25)	1.27 (0.67,2.42)	1.35 (0.67,2.70)	1.34 (0.67,2.68)
BMI 30+	1.25 (0.67,2.33)	1.43 (0.71,2.86)	1.47 (0.74,2.94)
BMI unknown	1.08 (0.58,1.98)	1.41 (0.72,2.75)	1.52 (0.78,2.96)
5-9 BNF chapters (*reference*: 0-4)	1.81 (1.19,2.73)	1.33 (0.84,2.11)	1.29 (0.82,2.05)
10+ BNF chapters	3.66 (2.21,6.08)	3.11 (1.75,5.55)	3.21 (1.81,5.69)
Multiple consultations in phase 1 (vs. single)	1.69 (1.17,2.44)	1.73 (1.14,2.63)	1.84 (1.22,2.79)
X-ray, phase 1[*] (vs. none)	1.45 (0.85,2.48)	1.28 (0.68,2.42)	1.20 (0.64,2.26)
Above the clinician median index consultation count (*reference*: at or below the median)	1.01 (0.50,2.04)	1.03 (0.47,2.29)	0.93 (0.43,2.03)

PPI prescription within 14 days of a clinical OA consultation, phase 1 (in those prescribed an oral NSAID)	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.44 (0.16,1.23)	0.56 (0.18,1.77)	-
Practice 3	0.98 (0.57,1.70)	0.88 (0.47,1.65)	-
Practice 4	0.72 (0.39,1.35)	0.77 (0.38,1.58)	-
Practice 5	2.31 (1.06,5.07)	2.19 (0.91,5.28)	-
Practice 6	0.88 (0.35,2.21)	0.86 (0.30,2.46)	-
Practice 7	2.50 (1.06,5.88)	2.90 (1.11,7.55)	-
Practice 8	1.01 (0.41,2.49)	0.92 (0.34,2.54)	-

[a]recorded relevant X-ray in phase 1 prior to the index clinical OA consultation

Estimates of associations (OR, 95%CI) between PPI prescription within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.3. Referral management

Appendix F.3.1. Physiotherapy referral

Physiotherapy referral within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	0.81 (0.59,1.11)	0.51 (0.35,0.75)	0.52 (0.35,0.76)
Sex (*reference*: female)	0.92 (0.68,1.24)	0.87 (0.64,1.18)	0.87 (0.64,1.19)
Age 65-74 (*reference*: 45-64)	0.78 (0.55,1.11)	0.85 (0.58,1.23)	0.84 (0.58,1.23)
Age 75-84	0.73 (0.48,1.09)	0.81 (0.52,1.28)	0.81 (0.51,1.27)
Age 85+	0.47 (0.24,0.93)	0.52 (0.25,1.07)	0.52 (0.25,1.06)
Hip (*reference*: knee)	0.97 (0.68,1.40)	0.99 (0.68,1.44)	0.99 (0.68,1.45)
Ankle/foot	0.21 (0.07,0.68)	0.24 (0.07,0.79)	0.24 (0.07,0.79)
Wrist/hand	0.47 (0.24,0.93)	0.48 (0.24,0.97)	0.49 (0.24,0.97)
Unspecified	0.20 (0.05,0.82)	0.30 (0.07,1.28)	0.30 (0.07,1.28)
Multisite	1.37 (0.91,2.08)	1.40 (0.88,2.21)	1.39 (0.88,2.20)
BMI 25 to <30 (reference: BMI <25)	0.71 (0.44,1.13)	0.68 (0.42,1.11)	0.68 (0.42,1.11)
BMI 30+	0.86 (0.55,1.33)	0.75 (0.47,1.20)	0.75 (0.47,1.19)
BMI unknown	0.78 (0.50,1.22)	0.70 (0.44,1.11)	0.70 (0.44,1.11)
5-9 BNF chapters (*reference*: 0-4)	0.83 (0.59,1.17)	0.82 (0.56,1.19)	0.83 (0.57,1.20)
10+ BNF chapters	0.81 (0.55,1.19)	0.79 (0.51,1.22)	0.80 (0.51,1.24)
Multiple consultations in phase 1 (vs. single)	2.86 (2.11,3.86)	2.75 (2.00,3.79)	2.75 (2.00,3.79)
X-ray, phase 1[a] (vs. none)	2.18 (1.44,3.31)	1.78 (1.14,2.79)	1.79 (1.14,2.81)
Above the clinician median index consultation count (*reference*: at or below the median)	1.04 (0.47,2.28)	1.23 (0.64,2.34)	1.22 (0.63,2.36)

Physiotherapy referral within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.62 (0.29,1.30)	0.75 (0.35,1.61)	-
Practice 3	0.18 (0.09,0.35)	0.16 (0.08,0.32)	-
Practice 4	0.01 (0.00,0.10)	0.02 (0.00,0.11)	-
Practice 5	0.00 (0.00,0.00)	0.00 (0.00,0.00)	-
Practice 6	0.52 (0.24,1.12)	0.48 (0.21,1.06)	-
Practice 7	0.46 (0.26,0.79)	0.40 (0.22,0.71)	-
Practice 8	0.09 (0.02,0.37)	0.07 (0.02,0.32)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between physiotherapy referral within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.3.2. All primary care referrals

All primary care team referrals within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	0.77 (0.56,1.05)	0.44 (0.31,0.64)	0.44 (0.30,0.64)
Sex (*reference*: female)	0.87 (0.65,1.17)	0.85 (0.63,1.14)	0.85 (0.63,1.16)
Age 65-74 (*reference*: 45-64)	0.77 (0.54,1.09)	0.81 (0.56,1.16)	0.80 (0.56,1.16)
Age 75-84	0.74 (0.50,1.10)	0.81 (0.53,1.24)	0.80 (0.51,1.24)
Age 85+	0.63 (0.35,1.15)	0.67 (0.36,1.26)	0.67 (0.35,1.27)
Hip (*reference*: knee)	0.97 (0.67,1.40)	0.96 (0.67,1.40)	0.97 (0.66,1.42)
Ankle/foot	0.25 (0.09,0.76)	0.29 (0.10,0.84)	0.29 (0.10,0.86)
Wrist/hand	0.88 (0.51,1.51)	0.94 (0.55,1.63)	0.95 (0.54,1.67)
Unspecified	0.69 (0.30,1.60)	1.12 (0.48,2.62)	1.12 (0.46,2.71)
Multisite	1.61 (1.08,2.40)	1.62 (1.05,2.51)	1.63 (1.04,2.54)
BMI 25 to <30 (reference: BMI <25)	0.69 (0.44,1.09)	0.70 (0.45,1.11)	0.69 (0.43,1.11)
BMI 30+	0.85 (0.56,1.30)	0.77 (0.50,1.19)	0.77 (0.49,1.20)
BMI unknown	0.67 (0.44,1.04)	0.64 (0.41,0.99)	0.63 (0.40,0.99)
5-9 BNF chapters (*reference*: 0-4)	0.86 (0.61,1.20)	0.85 (0.59,1.22)	0.85 (0.59,1.23)
10+ BNF chapters	0.99 (0.69,1.42)	0.94 (0.62,1.41)	0.95 (0.62,1.44)
Multiple consultations in phase 1 (vs. single)	3.18 (2.36,4.28)	3.07 (2.26,4.15)	3.12 (2.27,4.28)
X-ray, phase 1[a] (vs. none)	2.28 (1.52,3.43)	1.95 (1.27,3.01)	2.01 (1.29,3.12)
Above the clinician median index consultation count (*reference*: at or below the median)	1.13 (0.50,2.53)	1.03 (0.57,1.86)	1.05 (0.56,1.97)

All primary care team referrals within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.58 (0.28,1.22)	0.65 (0.30,1.44)	-
Practice 3	0.15 (0.08,0.31)	0.14 (0.07,0.28)	-
Practice 4	0.05 (0.02,0.14)	0.05 (0.02,0.15)	-
Practice 5	0.04 (0.01,0.14)	0.04 (0.01,0.14)	-
Practice 6	0.56 (0.27,1.17)	0.48 (0.22,1.06)	-
Practice 7	0.52 (0.31,0.87)	0.41 (0.23,0.72)	-
Practice 8	0.15 (0.05,0.46)	0.12 (0.04,0.38)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between any referral to a primary care team member within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix F.3.3. All secondary care referrals

All secondary care team (specialist) referrals within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.50 (1.24,1.82)	1.31 (1.04,1.64)	1.30 (1.04,1.63)
Sex (*reference*: female)	1.18 (0.98,1.42)	1.18 (0.97,1.44)	1.18 (0.97,1.44)
Age 65-74 (*reference*: 45-64)	0.99 (0.80,1.22)	0.92 (0.72,1.16)	0.91 (0.72,1.16)
Age 75-84	0.67 (0.51,0.87)	0.59 (0.44,0.79)	0.59 (0.44,0.79)
Age 85+	0.39 (0.25,0.63)	0.34 (0.20,0.55)	0.34 (0.21,0.56)
Hip (*reference*: knee)	0.94 (0.75,1.19)	1.01 (0.80,1.29)	1.02 (0.80,1.30)
Ankle/foot	0.29 (0.17,0.51)	0.36 (0.20,0.64)	0.35 (0.20,0.63)
Wrist/hand	0.39 (0.25,0.60)	0.48 (0.31,0.76)	0.48 (0.30,0.75)
Unspecified	0.24 (0.12,0.48)	0.23 (0.11,0.47)	0.24 (0.12,0.49)
Multisite	1.23 (0.92,1.64)	0.78 (0.57,1.07)	0.78 (0.57,1.07)
BMI 25 to <30 (reference: BMI <25)	1.49 (1.08,2.07)	1.35 (0.95,1.90)	1.35 (0.96,1.90)
BMI 30+	1.57 (1.14,2.16)	1.23 (0.88,1.73)	1.24 (0.89,1.74)
BMI unknown	1.29 (0.94,1.77)	1.15 (0.83,1.61)	1.15 (0.82,1.60)
5-9 BNF chapters (*reference*: 0-4)	1.10 (0.89,1.36)	1.04 (0.82,1.32)	1.04 (0.82,1.32)
10+ BNF chapters	0.93 (0.73,1.18)	0.98 (0.74,1.30)	0.98 (0.74,1.29)
Multiple consultations in phase 1 (vs. single)	3.46 (2.85,4.19)	3.24 (2.64,3.98)	3.22 (2.62,3.96)
X-ray, phase 1[a] (vs. none)	2.54 (1.89,3.43)	1.67 (1.19,2.33)	1.66 (1.19,2.31)
Above the clinician median index consultation count (*reference*: at or below the median)	1.18 (0.79,1.77)	1.25 (0.82,1.92)	1.26 (0.82,1.93)

All secondary care team (specialist) referrals within 14 days of a clinical OA consultation, phase 1	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.76 (0.43,1.33)	0.93 (0.52,1.65)	-
Practice 3	0.64 (0.44,0.93)	0.71 (0.48,1.04)	-
Practice 4	1.16 (0.83,1.61)	1.67 (1.18,2.36)	-
Practice 5	0.64 (0.43,0.95)	0.71 (0.47,1.07)	-
Practice 6	1.16 (0.69,1.93)	1.33 (0.78,2.27)	-
Practice 7	0.97 (0.67,1.40)	1.11 (0.75,1.65)	-
Practice 8	1.07 (0.65,1.77)	0.90 (0.53,1.53)	-

[a]**recorded relevant X-ray in phase 1 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between any referral to a secondary care speciality within 14 days of a consultation for clinical OA in phase one and the independent variables. Adjusted models include all independent variables.

Appendix G. Template-derived quality of care appendices (Chapter Six)

Appendix G.1. Template triggering patterns

The descriptive epidemiological characteristics of patients triggering versus not triggering the template are shown below. Non-triggering of the template occurred in all practices but was most frequent in practice 4 and to a lesser extent in practices 3 and 7. There was no substantial difference between males and females or between age bands.

	Template triggered *n* (%)	**Template not triggered *n* (%)**
Total	1730 (93.5)	121 (6.5)
Diagnosis		
Joint pain	95 (7.7)	1142 (92.3)
OA	26 (4.2)	588 (95.8)
Sex		
M	711 (93.4)	50 (6.6)
F	1019 (93.5)	71 (6.5)
Age band		
45-64	819 (92.9)	63 (7.1)
65-74	443 (94.3)	27 (5.7)
75-84	351 (94.1)	22 (5.9)
85+	117 (92.9)	9 (7.1)
Practice		
1	612 (97.9%)	13 (2.1%)
2	93 (98.9%)	1 (1.1%)
3	220 (89.8%)	25 (10.2%)
4	202 (81.8%)	45 (18.2%)
5	231 (93.5%)	16 (6.5%)
6	96 (99.0%)	1 (1.0%)
7	163 (90.6%)	17 (9.4%)
8	113 (97.4%)	3 (2.6%)

Characteristics and practices of patients triggering or not triggering the recording template in phase 2

Appendix G.2. Template-derived quality indicators - multilevel models

Appendix G.2.1. Pain assessment

	Two level models		Three level model
Pain assessment, phase 2	**Unadjusted**	**Adjusted**	**(adjusted)**
OA diagnosis (*reference:* joint pain)	1.33 (1.02,1.73)	1.26 (0.93,1.72)	1.26 (0.93,1.72)
Sex (*reference*: female)	1.12 (0.88,1.42)	1.09 (0.84,1.41)	1.09 (0.84,1.41)
Age 65-74 (*reference*: 45-64)	1.20 (0.90,1.61)	1.21 (0.87,1.67)	1.21 (0.87,1.67)
Age 75-84	1.33 (0.97,1.82)	1.46 (1.02,2.10)	1.46 (1.02,2.10)
Age 85+	1.07 (0.66,1.74)	1.35 (0.78,2.35)	1.35 (0.77,2.35)
Hip (*reference*: knee)	0.89 (0.66,1.22)	0.92 (0.66,1.27)	0.91 (0.66,1.27)
Ankle/foot	0.33 (0.20,0.52)	0.40 (0.24,0.66)	0.40 (0.24,0.66)
Wrist/hand	0.69 (0.45,1.06)	0.83 (0.53,1.29)	0.83 (0.53,1.30)
Unspecified	0.61 (0.37,1.02)	0.67 (0.39,1.15)	0.67 (0.39,1.16)
Multisite	1.08 (0.67,1.74)	0.66 (0.39,1.12)	0.66 (0.39,1.12)
BMI 25 to <30 (reference: BMI <25)	1.27 (0.89,1.81)	1.31 (0.90,1.91)	1.30 (0.90,1.90)
BMI 30+	1.24 (0.88,1.76)	1.35 (0.93,1.98)	1.35 (0.93,1.97)
BMI unknown	0.72 (0.50,1.06)	0.67 (0.45,1.00)	0.67 (0.45,1.00)
5-9 BNF chapters (*reference*: 0-4)	1.18 (0.87,1.60)	0.90 (0.64,1.26)	0.90 (0.65,1.27)
10+ BNF chapters	0.91 (0.68,1.22)	0.58 (0.41,0.84)	0.59 (0.41,0.84)
Multiple consultations in phase 2 (vs. single)	2.85 (2.16,3.76)	3.03 (2.24,4.11)	3.00 (2.21,4.08)
X-ray, phase 2[a] (vs. none)	1.26 (0.72,2.22)	1.04 (0.57,1.91)	1.07 (0.59,1.96)
Above the clinician median index consultation count (*reference*: at or below the median)	2.54 (1.14,5.64)	2.42 (1.11,5.27)	2.60 (1.15,5.86)

Pain assessment, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.34 (0.04,2.73)	1.49 (0.16,13.4)	-
Practice 3	0.54 (0.05,6.40)	0.48 (0.17,1.33)	-
Practice 4	0.04 (0.00,0.34)	0.22 (0.06,0.78)	-
Practice 5	0.12 (0.01,1.02)	0.08 (0.02,0.30)	-
Practice 6	0.18 (0.02,1.29)	1.20 (0.20,7.29)	-
Practice 7	0.85 (0.05,13.9)	0.70 (0.21,2.29)	-
Practice 8	0.47 (0.07,3.25)	2.31 (0.34,15.6)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between template-recorded pain assessment in phase two and the independent variables. Adjusted models include all independent variables.

Appendix G.2.2. Function assessment

Function assessment, phase 2	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.30 (1.00,1.68)	1.15 (0.85,1.56)	1.15 (0.85,1.56)
Sex (*reference*: female)	1.05 (0.83,1.33)	1.01 (0.78,1.30)	1.01 (0.78,1.30)
Age 65-74 (*reference*: 45-64)	1.29 (0.97,1.71)	1.32 (0.96,1.82)	1.32 (0.96,1.82)
Age 75-84	1.27 (0.94,1.72)	1.40 (0.98,1.99)	1.40 (0.98,2.00)
Age 85+	0.98 (0.61,1.57)	1.16 (0.67,1.98)	1.16 (0.67,1.99)
Hip (*reference*: knee)	0.84 (0.62,1.14)	0.84 (0.61,1.16)	0.84 (0.61,1.16)
Ankle/foot	0.30 (0.19,0.48)	0.35 (0.21,0.57)	0.35 (0.21,0.58)
Wrist/hand	0.76 (0.49,1.15)	0.90 (0.58,1.41)	0.91 (0.58,1.42)
Unspecified	0.51 (0.31,0.84)	0.56 (0.33,0.95)	0.56 (0.33,0.96)
Multisite	1.39 (0.86,2.27)	0.91 (0.53,1.55)	0.91 (0.53,1.56)
BMI 25 to <30 (reference: BMI <25)	1.26 (0.89,1.78)	1.31 (0.91,1.90)	1.31 (0.90,1.89)
BMI 30+	1.21 (0.86,1.70)	1.29 (0.89,1.88)	1.30 (0.89,1.88)
BMI unknown	0.65 (0.45,0.94)	0.59 (0.40,0.88)	0.59 (0.40,0.87)
5-9 BNF chapters (*reference*: 0-4)	1.00 (0.74,1.35)	0.72 (0.52,1.01)	0.73 (0.52,1.01)
10+ BNF chapters	0.91 (0.68,1.21)	0.58 (0.41,0.83)	0.58 (0.41,0.83)
Multiple consultations in phase 2 (vs. single)	3.04 (2.31,4.00)	3.11 (2.31,4.19)	3.11 (2.31,4.20)
X-ray, phase 2[a] (vs. none)	1.53 (0.87,2.67)	1.37 (0.75,2.49)	1.41 (0.77,2.57)
Above the clinician median index consultation count (*reference*: at or below the median)	3.30 (1.57,6.97)	3.28 (1.58,6.82)	3.58 (1.66,7.73)

Function assessment, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.40 (0.06,2.79)	1.94 (0.24,15.5)	-
Practice 3	0.72 (0.07,7.43)	0.48 (0.18,1.24)	-
Practice 4	0.05 (0.01,0.39)	0.29 (0.09,0.93)	-
Practice 5	0.15 (0.02,1.09)	0.11 (0.03,0.38)	-
Practice 6	0.17 (0.03,1.08)	1.66 (0.30,9.22)	-
Practice 7	1.21 (0.09,17.0)	0.86 (0.28,2.60)	-
Practice 8	0.48 (0.08,2.87)	2.40 (0.41,14.0)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between template-recorded function assessment in phase two and the independent variables. Adjusted models include all independent variables.

Appendix G.2.3. Weight record within 14 days of a clinical OA consultation

Weight record within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.05 (0.81,1.37)	0.95 (0.70,1.30)	0.97 (0.71,1.32)
Sex (*reference*: female)	1.20 (0.93,1.53)	1.27 (0.98,1.65)	1.27 (0.97,1.65)
Age 65-74 (*reference*: 45-64)	1.22 (0.91,1.63)	1.12 (0.82,1.55)	1.12 (0.81,1.55)
Age 75-84	0.90 (0.64,1.25)	0.90 (0.62,1.30)	0.89 (0.61,1.30)
Age 85+	0.53 (0.30,0.93)	0.58 (0.31,1.09)	0.58 (0.31,1.08)
Hip (*reference*: knee)	0.99 (0.72,1.36)	1.06 (0.75,1.48)	1.06 (0.75,1.49)
Ankle/foot	0.81 (0.49,1.35)	0.97 (0.57,1.65)	0.97 (0.56,1.65)
Wrist/hand	0.64 (0.40,1.05)	0.69 (0.41,1.14)	0.68 (0.41,1.13)
Unspecified	0.98 (0.58,1.65)	1.19 (0.68,2.08)	1.20 (0.69,2.11)
Multisite	1.36 (0.87,2.14)	1.34 (0.82,2.20)	1.34 (0.81,2.20)
BMI 25 to <30 (reference: BMI <25)	0.84 (0.59,1.18)	0.78 (0.54,1.11)	0.77 (0.53,1.10)
BMI 30+	1.17 (0.83,1.64)	1.09 (0.76,1.56)	1.08 (0.75,1.55)
BMI unknown	0.30 (0.18,0.48)	0.26 (0.16,0.42)	0.25 (0.15,0.42)
5-9 BNF chapters (*reference*: 0-4)	1.46 (1.07,1.99)	1.28 (0.90,1.80)	1.29 (0.91,1.83)
10+ BNF chapters	0.94 (0.69,1.29)	0.77 (0.53,1.12)	0.78 (0.53,1.13)
Multiple consultations in phase 2 (vs. single)	2.02 (1.55,2.62)	2.13 (1.61,2.82)	2.12 (1.60,2.82)
X-ray, phase 2[a] (vs. none)	1.22 (0.86,1.73)	0.88 (0.49,1.56)	0.89 (0.50,1.59)
Above the clinician median index consultation count (*reference*: at or below the median)	1.07 (0.63,1.81)	1.31 (0.69,2.46)	1.47 (0.76,2.86)

Weight record within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.54 (0.79,3.01)	1.94 (0.50,7.53)	-
Practice 3	1.84 (0.48,7.03)	0.35 (0.15,0.79)	-
Practice 4	0.32 (0.14,0.71)	0.95 (0.39,2.33)	-
Practice 5	0.97 (0.40,2.33)	0.30 (0.12,0.77)	-
Practice 6	0.34 (0.13,0.85)	0.93 (0.27,3.22)	-
Practice 7	1.21 (0.36,4.08)	0.22 (0.08,0.60)	-
Practice 8	0.26 (0.10,0.71)	1.52 (0.45,5.14)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between weight record within 14 days of a clinical OA consultation in phase two and the independent variables. Adjusted models include all independent variables.

Appendix G.2.4. Assessment of paracetamol use

	Two level models		Three level model
Paracetamol assessment, phase 2	**Unadjusted**	**Adjusted**	**(adjusted)**
OA diagnosis (*reference:* joint pain)	1.50 (1.18,1.92)	1.32 (0.99,1.76)	1.33 (1.00,1.77)
Sex (*reference*: female)	0.90 (0.72,1.13)	0.90 (0.71,1.14)	0.90 (0.71,1.15)
Age 65-74 (*reference*: 45-64)	1.69 (1.28,2.22)	1.63 (1.20,2.22)	1.65 (1.22,2.25)
Age 75-84	1.60 (1.19,2.15)	1.62 (1.15,2.27)	1.64 (1.16,2.30)
Age 85+	1.55 (0.98,2.45)	1.70 (1.02,2.83)	1.72 (1.03,2.88)
Hip (*reference*: knee)	1.07 (0.80,1.43)	1.03 (0.76,1.39)	1.03 (0.75,1.39)
Ankle/foot	0.37 (0.23,0.58)	0.44 (0.27,0.71)	0.44 (0.27,0.71)
Wrist/hand	0.61 (0.41,0.91)	0.72 (0.47,1.10)	0.71 (0.47,1.10)
Unspecified	0.64 (0.39,1.03)	0.61 (0.37,1.03)	0.62 (0.37,1.03)
Multisite	1.28 (0.81,2.00)	0.73 (0.44,1.19)	0.72 (0.44,1.19)
BMI 25 to <30 (reference: BMI <25)	1.09 (0.78,1.51)	1.09 (0.77,1.54)	1.08 (0.76,1.54)
BMI 30+	1.04 (0.75,1.44)	1.08 (0.76,1.54)	1.08 (0.76,1.54)
BMI unknown	0.62 (0.44,0.89)	0.61 (0.42,0.89)	0.60 (0.41,0.87)
5-9 BNF chapters (*reference*: 0-4)	1.48 (1.11,1.96)	1.08 (0.79,1.48)	1.09 (0.79,1.48)
10+ BNF chapters	1.23 (0.93,1.61)	0.72 (0.51,1.01)	0.72 (0.51,1.01)
Multiple consultations in phase 2 (vs. single)	2.90 (2.24,3.75)	2.98 (2.26,3.94)	3.02 (2.28,3.99)
X-ray, phase 2[a] (vs. none)	1.59 (0.93,2.71)	1.33 (0.76,2.35)	1.38 (0.78,2.43)
Above the clinician median index consultation count (*reference*: at or below the median)	3.37 (1.70,6.67)	3.18 (1.66,6.08)	3.41 (1.70,6.84)

Paracetamol assessment, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.73 (0.14,3.72)	0.60 (0.12,3.02)	-
Practice 3	0.35 (0.16,0.80)	0.48 (0.21,1.13)	-
Practice 4	0.40 (0.14,1.10)	0.36 (0.13,0.99)	-
Practice 5	0.15 (0.05,0.42)	0.16 (0.06,0.46)	-
Practice 6	2.42 (0.53,11.0)	2.87 (0.63,13.1)	-
Practice 7	0.66 (0.26,1.69)	0.69 (0.27,1.79)	-
Practice 8	3.35 (0.69,16.4)	3.40 (0.72,16.0)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between template-derived assessment of paracetamol use in phase two and the independent variables. Adjusted models include all independent variables.

Appendix G.2.5. Assessment of topical NSAID use

	Two level models		Three level model
Topical NSAID assessment, phase 2	**Unadjusted**	**Adjusted**	**(adjusted)**
OA diagnosis (*reference:* joint pain)	1.47 (1.14,1.90)	1.36 (1.01,1.83)	1.36 (1.01,1.83)
Sex (*reference*: female)	0.93 (0.73,1.17)	0.87 (0.68,1.12)	0.87 (0.68,1.12)
Age 65-74 (*reference*: 45-64)	1.37 (1.04,1.82)	1.38 (1.01,1.88)	1.38 (1.01,1.88)
Age 75-84	1.38 (1.02,1.87)	1.50 (1.06,2.13)	1.50 (1.06,2.13)
Age 85+	1.00 (0.62,1.62)	1.22 (0.71,2.10)	1.23 (0.72,2.10)
Hip (*reference*: knee)	0.69 (0.51,0.92)	0.68 (0.50,0.93)	0.68 (0.50,0.93)
Ankle/foot	0.46 (0.29,0.74)	0.56 (0.34,0.93)	0.56 (0.34,0.93)
Wrist/hand	0.71 (0.47,1.07)	0.86 (0.55,1.34)	0.86 (0.55,1.34)
Unspecified	0.44 (0.26,0.73)	0.43 (0.25,0.75)	0.43 (0.25,0.75)
Multisite	1.51 (0.95,2.41)	0.93 (0.56,1.54)	0.93 (0.56,1.55)
BMI 25 to <30 (reference: BMI <25)	1.50 (1.07,2.10)	1.48 (1.04,2.12)	1.47 (1.03,2.10)
BMI 30+	1.65 (1.18,2.31)	1.71 (1.19,2.45)	1.69 (1.18,2.43)
BMI unknown	0.88 (0.60,1.28)	0.83 (0.56,1.23)	0.82 (0.55,1.21)
5-9 BNF chapters (*reference*: 0-4)	1.36 (1.02,1.82)	1.01 (0.73,1.39)	1.01 (0.74,1.40)
10+ BNF chapters	1.12 (0.84,1.49)	0.69 (0.48,0.97)	0.69 (0.49,0.98)
Multiple consultations in phase 2 (vs. single)	2.73 (2.10,3.54)	2.75 (2.08,3.65)	2.75 (2.07,3.64)
X-ray, phase 2[a] (vs. none)	1.14 (0.67,1.92)	1.03 (0.58,1.82)	1.05 (0.60,1.86)
Above the clinician median index consultation count (*reference*: at or below the median)	3.14 (1.43,6.92)	2.87 (1.28,6.45)	3.25 (1.41,7.49)

Topical NSAID assessment, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.86 (0.11,6.47)	0.68 (0.08,5.54)	-
Practice 3	0.18 (0.06,0.49)	0.22 (0.07,0.65)	-
Practice 4	0.44 (0.13,1.50)	0.41 (0.11,1.47)	-
Practice 5	0.15 (0.04,0.53)	0.15 (0.04,0.57)	-
Practice 6	1.40 (0.26,7.56)	1.53 (0.25,9.27)	-
Practice 7	0.49 (0.15,1.53)	0.48 (0.14,1.60)	-
Practice 8	0.80 (0.13,4.77)	0.81 (0.12,5.45)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between template-derived assessment of topical NSAID use in phase two and the independent variables. Adjusted models include all independent variables.

Appendix G.2.6. Provision of education or information

Education provision, phase 2	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.68 (1.30,2.15)	1.60 (1.19,2.14)	1.61 (1.20,2.16)
Sex (*reference*: female)	1.04 (0.83,1.31)	1.01 (0.79,1.29)	1.01 (0.79,1.29)
Age 65-74 (*reference*: 45-64)	1.45 (1.10,1.91)	1.47 (1.08,2.01)	1.48 (1.09,2.01)
Age 75-84	1.25 (0.93,1.68)	1.30 (0.92,1.83)	1.30 (0.92,1.83)
Age 85+	1.15 (0.73,1.83)	1.36 (0.81,2.31)	1.36 (0.80,2.30)
Hip (*reference*: knee)	0.81 (0.61,1.09)	0.81 (0.59,1.10)	0.80 (0.59,1.09)
Ankle/foot	0.39 (0.24,0.62)	0.47 (0.29,0.78)	0.47 (0.29,0.78)
Wrist/hand	0.64 (0.42,0.97)	0.79 (0.51,1.22)	0.78 (0.50,1.22)
Unspecified	0.47 (0.28,0.78)	0.43 (0.25,0.74)	0.43 (0.25,0.74)
Multisite	1.31 (0.84,2.05)	0.74 (0.46,1.21)	0.73 (0.45,1.20)
BMI 25 to <30 (reference: BMI <25)	1.35 (0.97,1.89)	1.34 (0.94,1.90)	1.33 (0.94,1.90)
BMI 30+	1.37 (0.99,1.90)	1.39 (0.97,1.97)	1.39 (0.97,1.98)
BMI unknown	0.87 (0.60,1.25)	0.82 (0.56,1.21)	0.81 (0.55,1.20)
5-9 BNF chapters (*reference*: 0-4)	1.18 (0.88,1.57)	0.88 (0.64,1.20)	0.88 (0.64,1.21)
10+ BNF chapters	1.02 (0.77,1.35)	0.64 (0.45,0.90)	0.64 (0.45,0.90)
Multiple consultations in phase 2 (vs. single)	2.83 (2.19,3.66)	2.87 (2.18,3.77)	2.88 (2.18,3.82)
X-ray, phase 2[a] (vs. none)	1.34 (0.79,2.25)	1.06 (0.61,1.85)	1.07 (0.61,1.88)
Above the clinician median index consultation count (*reference*: at or below the median)	3.94 (1.80,8.62)	3.16 (1.54,6.48)	3.47 (1.62,7.43)

Education provision, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.63 (0.27,9.87)	1.40 (0.22,8.75)	-
Practice 3	0.12 (0.05,0.30)	0.14 (0.05,0.37)	-
Practice 4	0.39 (0.13,1.13)	0.36 (0.12,1.09)	-
Practice 5	0.10 (0.03,0.32)	0.10 (0.03,0.33)	-
Practice 6	1.80 (0.40,8.17)	2.14 (0.45,10.2)	-
Practice 7	0.38 (0.14,1.02)	0.39 (0.14,1.08)	-
Practice 8	1.89 (0.40,8.93)	2.07 (0.42,10.2)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between template-derived education provision achievement in phase 2 and the independent variables. Adjusted models include all independent variables.

Appendix G.2.7. Provision of exercise advice

	Two level models		Three level model
Exercise advice provision, phase 2	**Unadjusted**	**Adjusted**	**(adjusted)**
OA diagnosis (*reference:* joint pain)	1.66 (1.30,2.12)	1.56 (1.18,2.08)	1.58 (1.19,2.11)
Sex (*reference*: female)	1.08 (0.87,1.36)	1.04 (0.82,1.32)	1.04 (0.82,1.32)
Age 65-74 (*reference*: 45-64)	1.40 (1.06,1.84)	1.36 (1.01,1.83)	1.37 (1.01,1.85)
Age 75-84	1.09 (0.81,1.46)	1.06 (0.76,1.48)	1.06 (0.76,1.49)
Age 85+	0.71 (0.45,1.13)	0.74 (0.44,1.24)	0.74 (0.44,1.24)
Hip (*reference*: knee)	0.78 (0.59,1.04)	0.79 (0.58,1.06)	0.78 (0.58,1.06)
Ankle/foot	0.40 (0.25,0.64)	0.46 (0.28,0.74)	0.46 (0.28,0.75)
Wrist/hand	0.40 (0.26,0.61)	0.45 (0.29,0.69)	0.44 (0.28,0.69)
Unspecified	0.58 (0.35,0.94)	0.53 (0.32,0.89)	0.53 (0.31,0.89)
Multisite	1.33 (0.85,2.07)	0.90 (0.55,1.46)	0.89 (0.54,1.45)
BMI 25 to <30 (reference: BMI <25)	1.40 (1.01,1.95)	1.31 (0.93,1.85)	1.31 (0.93,1.85)
BMI 30+	1.41 (1.02,1.95)	1.29 (0.91,1.82)	1.30 (0.92,1.84)
BMI unknown	0.77 (0.54,1.11)	0.67 (0.46,0.98)	0.66 (0.45,0.97)
5-9 BNF chapters (*reference*: 0-4)	1.14 (0.86,1.50)	0.88 (0.64,1.20)	0.88 (0.65,1.20)
10+ BNF chapters	0.93 (0.71,1.23)	0.65 (0.47,0.91)	0.65 (0.47,0.92)
Multiple consultations in phase 2 (vs. single)	2.24 (1.75,2.87)	2.20 (1.69,2.86)	2.21 (1.70,2.89)
X-ray, phase 2[a] (vs. none)	1.45 (0.87,2.44)	1.15 (0.66,2.00)	1.17 (0.67,2.04)
Above the clinician median index consultation count (*reference*: at or below the median)	3.00 (1.50,6.03)	2.43 (1.28,4.59)	2.65 (1.34,5.25)

Exercise advice provision, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.57 (0.32,7.80)	1.40 (0.28,7.05)	-
Practice 3	0.18 (0.08,0.40)	0.22 (0.09,0.50)	-
Practice 4	0.40 (0.15,1.05)	0.40 (0.15,1.07)	-
Practice 5	0.14 (0.05,0.38)	0.14 (0.05,0.40)	-
Practice 6	1.94 (0.49,7.72)	2.17 (0.53,8.87)	-
Practice 7	0.46 (0.19,1.12)	0.49 (0.20,1.23)	-
Practice 8	2.49 (0.60,10.4)	2.61 (0.62,10.9)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between template-derived exercise advice provision achievement in phase 2 and the independent variables. Adjusted models include all independent variables.

Appendix G.2.8. Provision of weight loss advice

Weight loss advice provision, phase 2 [in those patients known to be overweight]	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.46 (1.06,2.01)	1.29 (0.91,1.85)	1.31 (0.92,1.88)
Sex (*reference*: female)	0.95 (0.71,1.28)	0.96 (0.70,1.30)	0.95 (0.70,1.29)
Age 65-74 (*reference*: 45-64)	1.12 (0.80,1.57)	1.21 (0.83,1.76)	1.21 (0.83,1.76)
Age 75-84	0.93 (0.63,1.38)	0.99 (0.63,1.53)	0.99 (0.63,1.54)
Age 85+	0.78 (0.38,1.61)	0.77 (0.36,1.68)	0.78 (0.36,1.69)
Hip (*reference*: knee)	0.79 (0.54,1.18)	0.78 (0.52,1.18)	0.79 (0.52,1.18)
Ankle/foot	0.46 (0.25,0.83)	0.54 (0.29,1.01)	0.53 (0.28,1.01)
Wrist/hand	0.46 (0.26,0.81)	0.52 (0.29,0.94)	0.51 (0.28,0.93)
Unspecified	0.63 (0.33,1.22)	0.60 (0.31,1.19)	0.61 (0.31,1.20)
Multisite	1.16 (0.65,2.07)	0.82 (0.44,1.51)	0.80 (0.43,1.49)
BMI 30+(reference: BMI 25.0-29.9)	1.32 (0.98,1.77)	1.29 (0.94,1.76)	1.30 (0.95,1.77)
5-9 BNF chapters (*reference*: 0-4)	0.94 (0.63,1.40)	0.79 (0.52,1.22)	0.80 (0.52,1.23)
10+ BNF chapters	0.91 (0.62,1.33)	0.71 (0.45,1.11)	0.72 (0.46,1.13)
Multiple consultations in phase 2 (vs. single)	2.02 (1.47,2.78)	2.00 (1.42,2.80)	1.99 (1.41,2.80)
X-ray, phase 2[a] (vs. none)	1.55 (0.80,3.01)	1.32 (0.66,2.66)	1.35 (0.67,2.73)
Above the clinician median index consultation count (*reference*: at or below the median)	2.67 (1.23,5.79)	1.99 (0.99,3.96)	2.13 (1.02,4.41)
Practice 2 (*reference*: Practice 1)	2.70 (0.62,11.8)	2.55 (0.53,12.4)	-
Practice 3	0.24 (0.10,0.56)	0.24 (0.10,0.60)	-
Practice 4	0.46 (0.18,1.16)	0.45 (0.17,1.19)	-
Practice 5	0.10 (0.04,0.29)	0.10 (0.03,0.30)	-
Practice 6	2.87 (0.77,10.6)	3.02 (0.76,12.0)	-
Practice 7	0.26 (0.11,0.66)	0.28 (0.11,0.75)	-
Practice 8	2.81 (0.75,10.6)	2.95 (0.73,11.9)	-

[a]recorded relevant X-ray in phase two prior to the index clinical OA consultation

Estimates of associations (OR, 95%CI) between template-derived weight loss advice provision achievement (in people known to be overweight) in phase 2 and the independent variables. Adjusted models include all independent variables.

Appendix G.2.9. Consideration of physiotherapy referral

	Two level models		Three level model
Physiotherapy consideration, phase 2	**Unadjusted**	**Adjusted**	**(adjusted)**
OA diagnosis (*reference:* joint pain)	1.38 (1.08,1.77)	1.34 (1.01,1.78)	1.36 (1.02,1.80)
Sex (*reference*: female)	1.09 (0.87,1.37)	1.12 (0.89,1.42)	1.12 (0.88,1.43)
Age 65-74 (*reference*: 45-64)	0.82 (0.62,1.07)	0.80 (0.59,1.07)	0.79 (0.58,1.07)
Age 75-84	1.08 (0.80,1.45)	1.06 (0.76,1.48)	1.05 (0.75,1.47)
Age 85+	1.01 (0.64,1.60)	1.02 (0.62,1.69)	1.02 (0.62,1.69)
Hip (*reference*: knee)	0.95 (0.71,1.27)	0.97 (0.72,1.30)	0.96 (0.71,1.30)
Ankle/foot	0.63 (0.39,1.02)	0.79 (0.48,1.31)	0.79 (0.47,1.30)
Wrist/hand	0.93 (0.61,1.41)	1.10 (0.71,1.69)	1.09 (0.71,1.69)
Unspecified	0.49 (0.29,0.85)	0.50 (0.28,0.88)	0.49 (0.28,0.88)
Multisite	1.49 (0.98,2.25)	1.05 (0.67,1.65)	1.05 (0.67,1.65)
BMI 25 to <30 (reference: BMI <25)	0.97 (0.70,1.34)	0.93 (0.66,1.31)	0.93 (0.66,1.30)
BMI 30+	1.08 (0.78,1.49)	1.03 (0.73,1.46)	1.04 (0.73,1.46)
BMI unknown	0.86 (0.60,1.24)	0.76 (0.52,1.12)	0.76 (0.52,1.11)
5-9 BNF chapters (*reference*: 0-4)	1.04 (0.78,1.38)	0.96 (0.70,1.31)	0.96 (0.71,1.31)
10+ BNF chapters	1.02 (0.77,1.35)	0.86 (0.62,1.21)	0.87 (0.62,1.21)
Multiple consultations in phase 2 (vs. single)	2.21 (1.74,2.82)	2.20 (1.71,2.84)	2.22 (1.72,2.87)
X-ray, phase 2[a] (vs. none)	1.49 (0.92,2.42)	1.26 (0.76,2.12)	1.29 (0.77,2.17)
Above the clinician median index consultation count (*reference*: at or below the median)	2.77 (1.44,5.34)	2.49 (1.32,4.73)	2.70 (1.38,5.25)

Physiotherapy consideration, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.39 (0.32,5.98)	1.20 (0.27,5.32)	-
Practice 3	0.29 (0.14,0.62)	0.35 (0.16,0.76)	-
Practice 4	0.40 (0.16,0.98)	0.42 (0.17,1.06)	-
Practice 5	0.10 (0.04,0.28)	0.10 (0.03,0.29)	-
Practice 6	0.83 (0.24,2.88)	0.88 (0.24,3.28)	-
Practice 7	0.53 (0.23,1.23)	0.56 (0.23,1.34)	-
Practice 8	0.58 (0.16,2.04)	0.53 (0.14,2.02)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between template-derived physiotherapy consideration achievement in phase 2 and the independent variables. Adjusted models include all independent variables.

Appendix G.2.10. Achievement of all eight template-derived indicators

Achievement of all eight template-derived indicators, phase 2	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.98 (1.47,2.67)	1.84 (1.30,2.60)	1.86 (1.31,2.65)
Sex (*reference*: female)	1.07 (0.81,1.41)	1.13 (0.84,1.51)	1.13 (0.84,1.52)
Age 65-74 (*reference*: 45-64)	1.01 (0.72,1.42)	0.94 (0.65,1.35)	0.93 (0.64,1.36)
Age 75-84	1.29 (0.89,1.86)	1.15 (0.76,1.74)	1.16 (0.76,1.76)
Age 85+	0.95 (0.53,1.68)	0.78 (0.41,1.47)	0.77 (0.40,1.48)
Hip (*reference*: knee)	0.93 (0.64,1.33)	0.94 (0.65,1.38)	0.94 (0.64,1.38)
Ankle/foot	0.76 (0.42,1.37)	0.96 (0.51,1.79)	0.96 (0.51,1.80)
Wrist/hand	0.75 (0.44,1.29)	0.92 (0.52,1.61)	0.91 (0.52,1.61)
Unspecified	0.81 (0.42,1.54)	0.69 (0.35,1.37)	0.70 (0.35,1.39)
Multisite	1.75 (1.10,2.80)	1.11 (0.67,1.83)	1.10 (0.67,1.83)
BMI 25 to <30 (reference: BMI <25)	0.90 (0.61,1.34)	0.81 (0.53,1.22)	0.80 (0.53,1.22)
BMI 30+	0.96 (0.64,1.43)	0.83 (0.54,1.27)	0.83 (0.54,1.27)
BMI unknown	0.90 (0.57,1.42)	0.80 (0.50,1.29)	0.80 (0.49,1.28)
5-9 BNF chapters (*reference*: 0-4)	1.07 (0.75,1.53)	0.95 (0.64,1.40)	0.95 (0.64,1.41)
10+ BNF chapters	1.17 (0.83,1.66)	0.95 (0.62,1.44)	0.95 (0.62,1.46)
Multiple consultations in phase 2 (vs. single)	2.21 (1.64,2.98)	2.14 (1.56,2.92)	2.17 (1.58,2.98)
X-ray, phase 2[a] (vs. none)	1.69 (0.98,2.92)	1.28 (0.72,2.29)	1.31 (0.73,2.35)
Above the clinician median index consultation count (*reference*: at or below the median)	2.98 (1.21,7.32)	2.37 (1.04,5.42)	2.55 (1.05,6.20)

Achievement of all eight template-derived indicators, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.37 (0.07,1.88)	0.32 (0.06,1.71)	-
Practice 3	0.17 (0.06,0.45)	0.20 (0.07,0.56)	-
Practice 4	0.32 (0.11,0.90)	0.34 (0.12,1.01)	-
Practice 5	0.10 (0.03,0.35)	0.11 (0.03,0.39)	-
Practice 6	1.45 (0.38,5.51)	1.87 (0.45,7.77)	-
Practice 7	0.22 (0.07,0.65)	0.25 (0.08,0.79)	-
Practice 8	0.73 (0.18,2.93)	0.65 (0.15,2.93)	-

[a]**recorded relevant X-ray in phase 2 prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between all eight template-derived indicator (excludes weight measurement) achievement in phase 2 and the independent variables. Adjusted models include all independent variables.

Appendix G.2.11. Achievement of all eight template indicators in people referred on

Achievement of all eight template indicators in people referred for further care, phase 2	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.84 (0.93,3.65)	1.47 (0.63,3.43)	1.46 (0.63,3.39)
Sex (*reference*: female)	0.96 (0.51,1.82)	0.77 (0.37,1.59)	0.79 (0.38,1.64)
Age 65-74 (*reference*: 45-64)	0.95 (0.44,2.06)	0.81 (0.32,2.05)	0.79 (0.31,2.00)
Age 75-84	1.16 (0.45,2.99)	0.80 (0.26,2.47)	0.78 (0.25,2.43)
Age 85+	2.28 (0.42,12.28)	1.16 (0.14,9.71)	1.22 (0.14,10.4)
Hip (*reference*: knee)	1.09 (0.47,2.54)	0.94 (0.37,2.40)	0.92 (0.36,2.36)
Ankle/foot	0.98 (0.13,7.36)	1.15 (0.13,10.13)	1.34 (0.15,11.9)
Wrist/hand	0.42 (0.07,2.53)	0.39 (0.05,2.88)	0.38 (0.05,2.76)
Unspecified	0.40 (0.03,5.13)	1.13 (0.07,17.16)	1.04 (0.07,15.2)
Multisite	2.19 (0.78,6.16)	1.53 (0.46,5.10)	1.31 (0.40,4.34)
BMI 25 to <30 (reference: BMI <25)	1.51 (0.56,4.11)	1.09 (0.35,3.39)	1.16 (0.37,3.64)
BMI 30+	1.02 (0.36,2.87)	0.87 (0.27,2.77)	0.91 (0.29,2.92)
BMI unknown	0.78 (0.26,2.39)	0.70 (0.21,2.40)	0.71 (0.21,2.45)
5-9 BNF chapters (*reference*: 0-4)	1.14 (0.52,2.48)	0.88 (0.35,2.20)	0.89 (0.36,2.22)
10+ BNF chapters	1.60 (0.72,3.56)	1.19 (0.39,3.68)	1.18 (0.38,3.63)
Multiple consultations in phase 2 (vs. single)	4.40 (2.14,9.06)	4.45 (2.04,9.70)	4.38 (1.96,9.81)
X-ray, phase 2[a] (vs. none)	2.73 (0.96,7.74)	2.49 (0.77,8.05)	2.48 (0.75,8.24)
Above the clinician median index consultation count (*reference*: at or below the median)	3.54 (0.56,22.4)	5.02 (0.88,28.5)	4.95 (0.76,32.1)

Achievement of all eight template indicators in people referred for further care, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.14 (0.00,4.64)	0.07 (0.00,4.33)	-
Practice 3	0.29 (0.06,1.37)	0.30 (0.05,1.85)	-
Practice 4	0.05 (0.00,0.89)	0.04 (0.00,1.09)	-
Practice 5	0.15 (0.01,1.87)	0.08 (0.00,1.49)	-
Practice 6	7.19 (1.05,49.4)	7.77 (0.85,71.1)	-
Practice 7	0.19 (0.03,1.03)	0.18 (0.03,1.34)	-
Practice 8	2.09 (0.25,17.7)	1.51 (0.13,17.0)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95%CI) between all eight template-derived indicators (excludes weight measurement) achievement in phase 2 (in people referred for further care) and the independent variables. Adjusted models include all independent variables.

Appendix G.3. Pharmacological management - multilevel models

Appendix G.3.1. Recorded paracetamol prescription

Recorded paracetamol prescription within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.89 (1.44,2.48)	1.39 (1.00,1.92)	1.40 (1.01,1.93)
Sex (*reference*: female)	0.78 (0.60,1.02)	0.87 (0.66,1.16)	0.88 (0.66,1.18)
Age 65-74 (*reference*: 45-64)	2.09 (1.48,2.94)	1.54 (1.07,2.22)	1.55 (1.08,2.24)
Age 75-84	3.55 (2.53,4.99)	2.50 (1.72,3.63)	2.52 (1.73,3.68)
Age 85+	5.13 (3.22,8.17)	3.48 (2.06,5.87)	3.54 (2.09,5.99)
Hip (*reference*: knee)	1.23 (0.89,1.70)	1.16 (0.83,1.64)	1.15 (0.81,1.63)
Ankle/foot	0.60 (0.33,1.09)	0.77 (0.41,1.44)	0.75 (0.40,1.42)
Wrist/hand	0.59 (0.34,1.03)	0.71 (0.40,1.27)	0.70 (0.39,1.26)
Unspecified	1.10 (0.63,1.91)	0.90 (0.49,1.65)	0.87 (0.47,1.60)
Multisite	1.55 (0.98,2.47)	0.83 (0.49,1.39)	0.83 (0.49,1.40)
BMI 25 to <30 (reference: BMI <25)	0.84 (0.57,1.23)	0.74 (0.49,1.11)	0.74 (0.49,1.13)
BMI 30+	1.16 (0.80,1.69)	1.20 (0.80,1.80)	1.19 (0.79,1.79)
BMI unknown	0.79 (0.51,1.21)	0.87 (0.55,1.37)	0.87 (0.54,1.38)
5-9 BNF chapters (*reference*: 0-4)	3.59 (2.30,5.62)	2.72 (1.70,4.35)	2.72 (1.69,4.38)
10+ BNF chapters	5.75 (3.74,8.84)	3.32 (2.06,5.35)	3.32 (2.05,5.38)
Multiple consultations in phase 2 (vs. single)	2.06 (1.58,2.69)	1.98 (1.48,2.64)	1.98 (1.48,2.64)
X-ray, phase 2[a] (vs. none)	1.29 (0.74,2.25)	1.17 (0.63,2.16)	1.15 (0.62,2.13)
Above the clinician median index consultation count (*reference*: at or below the median)	0.99 (0.62,1.60)	0.97 (0.59,1.59)	0.92 (0.55,1.56)

Recorded paracetamol prescription within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.88 (0.40,1.96)	0.92 (0.37,2.28)	-
Practice 3	1.41 (0.88,2.26)	1.94 (1.13,3.31)	-
Practice 4	0.62 (0.35,1.11)	0.65 (0.34,1.24)	-
Practice 5	1.02 (0.60,1.72)	1.21 (0.66,2.22)	-
Practice 6	1.63 (0.79,3.34)	2.31 (1.01,5.27)	-
Practice 7	0.44 (0.23,0.87)	0.50 (0.24,1.05)	-
Practice 8	0.58 (0.26,1.31)	0.49 (0.20,1.20)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95% CI) between recorded prescription for paracetamol within 14 days of a clinical OA consultation in phase two and the independent variables. Adjusted models include all independent variables.

Appendix G.3.2. Recorded topical NSAID prescription

Recorded topical NSAID prescription within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
OA diagnosis (*reference:* joint pain)	1.37 (1.06,1.76)	0.92 (0.68,1.26)	0.94 (0.69,1.28)
Sex (*reference*: female)	0.76 (0.60,0.96)	0.80 (0.62,1.03)	0.79 (0.61,1.02)
Age 65-74 (*reference*: 45-64)	2.35 (1.73,3.20)	2.08 (1.46,2.96)	1.74 (1.27,2.39)
Age 75-84	3.79 (2.42,5.93)	3.46 (2.08,5.77)	2.02 (1.42,2.87)
Age 85+	1.37 (1.06,1.76)	0.92 (0.68,1.26)	3.34 (2.01,5.55)
Hip (*reference*: knee)	0.43 (0.31,0.61)	0.36 (0.25,0.52)	0.37 (0.26,0.53)
Ankle/foot	0.37 (0.21,0.65)	0.38 (0.21,0.68)	0.39 (0.22,0.70)
Wrist/hand	1.00 (0.67,1.50)	1.16 (0.76,1.79)	1.15 (0.75,1.77)
Unspecified	0.73 (0.44,1.22)	0.67 (0.38,1.15)	0.66 (0.38,1.15)
Multisite	1.32 (0.87,2.02)	0.93 (0.58,1.48)	0.90 (0.56,1.45)
BMI 25 to <30 (reference: BMI <25)	1.09 (0.77,1.53)	1.00 (0.70,1.45)	1.00 (0.69,1.44)
BMI 30+	1.28 (0.92,1.80)	1.25 (0.87,1.81)	1.24 (0.86,1.79)
BMI unknown	0.67 (0.44,1.00)	0.75 (0.49,1.16)	0.73 (0.48,1.13)
5-9 BNF chapters (*reference*: 0-4)	2.46 (1.75,3.46)	1.89 (1.31,2.72)	1.89 (1.31,2.73)
10+ BNF chapters	3.49 (2.51,4.86)	2.17 (1.48,3.18)	2.19 (1.49,3.21)
Multiple consultations in phase 2 (vs. single)	1.92 (1.50,2.45)	1.84 (1.41,2.40)	1.83 (1.40,2.39)
X-ray, phase 2[a] (vs. none)	0.96 (0.56,1.64)	1.10 (0.60,2.00)	1.08 (0.60,1.96)
Above the clinician median index consultation count (*reference*: at or below the median)	1.39 (0.82,2.33)	1.23 (0.70,2.14)	1.38 (0.78,2.42)

Recorded topical NSAID prescription within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	1.69 (0.59,4.85)	1.82 (0.56,5.90)	-
Practice 3	0.42 (0.22,0.82)	0.45 (0.22,0.93)	-
Practice 4	0.95 (0.47,1.91)	0.95 (0.44,2.06)	-
Practice 5	1.21 (0.62,2.36)	1.42 (0.67,3.01)	-
Practice 6	1.69 (0.65,4.40)	1.70 (0.59,4.95)	-
Practice 7	0.80 (0.40,1.61)	0.83 (0.38,1.81)	-
Practice 8	1.00 (0.37,2.68)	0.92 (0.31,2.75)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95% CI) between recorded prescription for topical NSAIDs within 14 days of a clinical OA consultation in phase two and the independent variables. Adjusted models include all independent variables.

Appendix G.4. Referral management - multilevel models

Appendix G.4.1. Recorded physiotherapy referral

Recorded physiotherapy referral within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model
	Unadjusted	Adjusted	(adjusted)
OA diagnosis (*reference:* joint pain)	1.02 (0.65,1.58)	1.08 (0.65,1.79)	1.04 (0.62,1.76)
Sex (*reference*: female)	1.59 (1.06,2.36)	1.46 (0.97,2.19)	1.47 (0.96,2.25)
Age 65-74 (*reference*: 45-64)	0.69 (0.41,1.16)	0.76 (0.44,1.30)	0.75 (0.43,1.32)
Age 75-84	1.01 (0.60,1.71)	1.06 (0.59,1.87)	1.07 (0.59,1.93)
Age 85+	0.80 (0.35,1.83)	0.87 (0.36,2.12)	0.88 (0.35,2.21)
Hip (*reference*: knee)	0.74 (0.45,1.21)	0.75 (0.45,1.24)	0.74 (0.44,1.25)
Ankle/foot	0.18 (0.04,0.81)	0.23 (0.05,1.03)	0.22 (0.05,1.09)
Wrist/hand	0.20 (0.06,0.69)	0.23 (0.07,0.80)	0.23 (0.06,0.83)
Unspecified	0.20 (0.04,0.91)	0.26 (0.06,1.20)	0.26 (0.05,1.27)
Multisite	0.56 (0.26,1.22)	0.45 (0.20,1.02)	0.45 (0.20,1.05)
BMI 25 to <30 (reference: BMI <25)	1.20 (0.67,2.17)	1.00 (0.55,1.82)	1.02 (0.54,1.90)
BMI 30+	1.15 (0.63,2.08)	0.99 (0.54,1.84)	1.00 (0.53,1.90)
BMI unknown	0.95 (0.48,1.86)	0.78 (0.40,1.55)	0.77 (0.38,1.57)
5-9 BNF chapters (*reference*: 0-4)	0.97 (0.59,1.60)	0.91 (0.54,1.54)	0.91 (0.53,1.58)
10+ BNF chapters	0.86 (0.52,1.40)	0.86 (0.47,1.54)	0.85 (0.46,1.57)
Multiple consultations in phase 2 (vs. single)	2.50 (1.68,3.74)	2.35 (1.56,3.53)	2.39 (1.56,3.66)
X-ray, phase 2[a] (vs. none)	2.24 (1.11,4.52)	1.87 (0.90,3.89)	1.98 (0.93,4.21)
Above the clinician median index consultation count (*reference*: at or below the median)	0.99 (0.45,2.20)	1.00 (0.49,2.03)	0.98 (0.46,2.10)

Recorded physiotherapy referral within 14 days of a clinical OA consultation, phase 2	Two level models		Three level model (adjusted)
	Unadjusted	Adjusted	
Practice 2 (*reference*: Practice 1)	0.44 (0.13,1.49)	0.44 (0.12,1.57)	-
Practice 3	0.48 (0.23,1.02)	0.49 (0.22,1.06)	-
Practice 4	0.12 (0.03,0.44)	0.13 (0.03,0.51)	-
Practice 5	0.04 (0.00,0.30)	0.04 (0.00,0.32)	-
Practice 6	1.67 (0.67,4.18)	1.59 (0.59,4.27)	-
Practice 7	0.83 (0.39,1.77)	0.78 (0.35,1.74)	-
Practice 8	0.22 (0.05,0.90)	0.19 (0.04,0.83)	-

[a]**recorded relevant X-ray in phase two prior to the index clinical OA consultation**

Estimates of associations (OR, 95% CI) between recorded referral to physiotherapy within 14 days of a clinical OA consultation in phase two and the independent variables. Adjusted models include all independent variables.

Appendix H. Assessment of the effects of the model OA consultation (MOSAICS cluster trial) appendices (Chapter Eight)

Appendix H.1. Adjusted odds of specified outcomes in intervention practices compared to control, trial period

Outcome measure	OR (intervention vs. control arm) (95% CI)		
	Unadjusted two-level model[d]	Adjusted two-level model[d]	Adjusted three-level model[e]
Assessment			
Pain assessment	↑ **2.69 (1.39,5.20)**	1.34 (0.54,3.32)	1.34 (0.54,3.32)
Function assessment	↑ **2.71 (1.41,5.22)**	1.14 (0.46,2.82)	1.14 (0.46,2.82)
Weight record	↑ **1.78 (1.07,2.97)**	1.15 (0.66,2.03)	1.54 (0.64,3.71)
X-ray recorded	↑ **3.43 (1.68,7.00)**	0.43 (0.09,1.99)	0.65 (0.03,12.5)
Core interventions			
OA information provision	↑ **2.97 (1.52,5.80)**	1.31 (0.56,3.03)	1.31 (0.56,3.03)
Written OA information	↑ **25.7 (8.89,74.4)**	↑ **23.2 (7.10,75.8)**	↑ **29.5 (7.41,118)**
Exercise advice provision	↑ **2.63 (1.45,4.77)**	1.50 (0.68,3.29)	1.50 (0.68,3.29)
Written exercise advice	↑ **27.6 (8.57,89.1)**	↑ **21.4 (6.69,68.8)**	↑ **25.5 (7.10,91.7)**
Weight loss advice provision[a]	↑ **3.08 (1.52,6.24)**	1.39 (0.66,2.96)	1.39 (0.66,2.96)
Written weight loss advice[ab]	↑ **24.8 (5.47,113)**	↑ **23.5 (4.61,120)**	↑ **28.8 (4.58,181)**
Non-pharmacological management			
Consideration of physiotherapy referral[b]	↑ **2.17 (1.18,3.97)**	1.41 (0.56,3.57)	1.41 (0.56,3.56)
Physiotherapy referral made	↑ **7.66 (3.07,19.1)**	↑ **5.50 (2.13,14.2)**	↑ **5.50 (2.13,14.2)**

Outcome measure	OR (intervention vs. control arm) (95% CI)		
	Unadjusted two-level model[d]	Adjusted two-level model[d]	Adjusted three-level model[e]
Pharmacological management			
Consideration of paracetamol use	↑ **2.69 (1.49,4.86)**	1.42 (0.67,3.04)	1.42 (0.67,3.04)
Paracetamol prescribed	↑ **1.77 (1.32,2.39)**	↑ **1.58 (1.14,2.20)**	↑ **1.58 (1.14,2.20)**
Consideration of topical NSAID use	↑ **2.00 (1.14,3.53)**	0.91 (0.43,1.92)	0.91 (0.43,1.92)
Topical NSAID prescribed	↑ **1.46 (1.01,2.10)**	1.13 (0.75,1.70)	1.13 (0.75,1.70)
Oral NSAID prescribed	1.26 (0.99,1.60)	0.79 (0.53,1.16)	0.79 (0.53,1.16)
Oral NSAID prescribed in the presence of a relative comorbid contraindication	0.74 (0.41,1.33)	0.74 (0.42,1.32)	0.74 (0.42,1.31)
Gastroprotection prescribed (PPI)[c]	0.90 (0.60,1.35)	0.96 (0.50,1.82)	0.80 (0.35,1.81)
Opioid prescribed	1.05 (0.62,1.78)	0.85 (0.66,1.09)	0.85 (0.66,1.09)

Adjusted for OA or joint pain code, sex, age band, site of disease, BMI status, morbidity load (BNF chapter count), multiple clinical OA consultation, staff member index consultation count dichotomy, and practice pre-trial achievement; [a]in those known to be overweight at time of index consultation; [b]PQL1 model; [c]denominator: those prescribed oral NSAIDs; [d]patients within clinicians; [e]patients within clinicians within practices

Intention-to-treat analysis: comparison of two-level unadjusted and adjusted model, and three-level model OR.

Appendix H.2. Adjusted odds of specified outcomes in intervention practices compared to control, restriction to patients triggering template

Outcome measure	Number of patients with outcome *n* (%)		Odds of outcome (intervention vs. control arm) (95% CI)			
	Intervention arm	Control arm	Unadjusted two-level model[d]	Adjusted two-level model[d]	Adjusted three-level model[e]	Adjusted two-level model[d] in new consulters[f] only (*n*=1205)
Total	1053	751				
Assessment						
Pain assessment	612 (58.1)	317 (42.2)	2.57 (1.33,4.97)	1.37 (0.55,3.45)	1.38 (0.54,3.51)	0.79 (0.29,2.14)
Function assessment	606 (57.5)	307 (40.9)	2.60 (1.35,5.02)	1.17 (0.46,2.93)	1.17 (0.46,2.93)	0.68 (0.24,1.88)
Weight record	301 (28.6)	136 (18.1)	1.73 (1.04,2.89)	1.17 (0.65,2.08)	1.58 (0.64,3.89)	0.92 (0.52,1.65)
X-ray recorded	154 (14.6)	43 (5.7)	3.32 (1.57,7.04)	0.38 (0.08,1.80)	0.58 (0.03,12.8)	0.30 (0.07,1.36)
Core interventions						
OA information provision	549 (52.1)	267 (35.6)	2.84 (1.47,5.47)	1.34 (0.58,3.13)	1.34 (0.58,3.13)	0.78 (0.32,1.90)
Written OA information	292 (27.7)	12 (1.6)	25.0 (8.75,71.5)	**↑ 24.2 (7.33,80.1)**	**↑ 32.3 (7.70,135)**	**↑ 26.5 (7.19,97.8)**
Exercise advice provision	522 (49.6)	245 (32.6)	2.51 (1.40,4.50)	1.56 (0.71,3.46)	1.56 (0.71,3.46)	0.93 (0.43,2.02)
Written exercise advice	230 (21.8)	7 (0.9)	26.8 (8.42,85.2)	**↑ 21.9 (6.79,70.4)**	**↑ 26.9 (7.25,99.6)**	**↑ 11.9 (4.19,33.8)**
Weight loss advice provision[a]	327 (47.3)	130 (29.0)	2.77 (1.37,5.60)	1.37 (0.64,2.94)	1.37 (0.64,2.94)	0.93 (0.42,2.06)
Written weight loss advice[ab]	99 (14.3)	2 (0.4)	30.2 (4.16,219)	**↑ 34.3 (4.05,291)**	**↑ 28.6 (4.61,178)**	**↑ 13.1 (2.34,73.5)**
Non-pharmacological management						
Consideration of physiotherapy referral	94 (8.9)	65 (8.7)	2.04 (1.11,3.73)	1.41 (0.55,3.59)	1.41 (0.55,3.58)	0.98 (0.38,2.51)
Physiotherapy referral made	107 (10.2)	19 (2.5)	7.09 (2.88,17.4)	**↑ 5.44 (2.15,13.8)**	**↑ 5.44 (2.14,13.8)**	**↑ 5.78 (1.77,18.9)**

(con't)

Outcome measure	Number of patients with outcome *n* (%)		Odds of outcome (intervention vs. control arm) (95% CI)			
	Intervention arm	Control arm	Unadjusted two-level model[d]	Adjusted two-level model[d]	Adjusted three-level model[e]	Adjusted two-level model[d] in new consulters[f] only (*n*=1205)
Pharmacological management						
Consideration of paracetamol use	549 (52.1)	282 (37.5)	2.54 (1.41,4.56)	1.43 (0.67,3.07)	1.43 (0.67,3.07)	0.97 (0.43,2.21)
Paracetamol prescribed	236 (22.4)	107 (14.2)	1.81 (1.34,2.45)	↑ **1.69 (1.20,2.37)**	↑ **1.69 (1.20,2.37)**	↑ **1.78 (1.13,2.78)**
Consideration of topical NSAID use	496 (47.1)	274 (36.5)	1.88 (1.07,3.31)	0.93 (0.44,1.99)	0.93 (0.44,1.99)	0.55 (0.24,1.25)
Topical NSAID prescribed	312 (29.6)	167 (22.2)	1.45 (1.00,2.10)	1.15 (0.77,1.72)	1.15 (0.77,1.72)	1.01 (0.62,1.64)
Oral NSAID prescribed	173 (16.4)	124 (16.5)	0.88 (0.58,1.35)	0.78 (0.52,1.16)	0.78 (0.52,1.16)	1.07 (0.67,1.71)
Oral NSAID prescribed in the presence of a relative comorbid contraindication	52 (12.4)	40 (15.7)	0.75 (0.41,1.37)	0.75 (0.41,1.35)	0.75 (0.41,1.35)	0.99 (0.50,1.93)
Gastroprotection prescribed (PPI)[c]	68 (39.3)	46 (37.1)	1.02 (0.59,1.78)	1.01 (0.52,1.97)	0.82 (0.35,1.96)	1.69 (0.48,5.95)
Opioid prescribed	353 (33.5)	208 (27.7)	1.27 (0.99,1.63)	0.88 (0.68,1.14)	0.88 (0.68,1.14)	0.91 (0.66,1.24)

Adjusted for OA or joint pain code, sex, age band, site of disease, BMI status, morbidity load (BNF chapter count), multiple clinical OA consultation, staff member index consultation count dichotomy, and practice pre-trial achievement [a]in those known to be overweight at time of index consultation [b]PQL1 model; [c]denominator: those prescribed oral NSAIDs; [d]patients within clinicians [e]patients within clinicians within practices [f]new consulters defined as first clinical OA consultation since the introduction of the template and with at least 365 days since any previous OA or joint pain consultation

Odds of outcome in intervention practices compared to control, trial period (95% CI) in patients for whom the template was recorded as triggering.

Appendix H.3. Adjusted odds of specified outcomes in intervention practices compared to control, restriction to patients with at least one template entry

Outcome measure	Number of patients with outcome *n* (%)		Odds of outcome (intervention vs. control arm) (95% CI)		
	Intervention arm	Control arm	Unadjusted two-level model[f]	Adjusted two-level model[f]	Adjusted three-level model[g]
Total	630	328			
Assessment					
Pain assessment[ab]	612 (97.1)	317 (96.6)	1.60 (0.54,4.68)	1.37 (0.33,5.73)	1.37 (0.33,5.73)
Function assessment[ab]	606 (96.2)	307 (93.6)	1.97 (0.70,5.50)	1.97 (0.70,5.50)	1.85 (0.71,4.83)
Weight record	275 (43.7)	91 (27.7)	1.78 (0.97,3.27)	1.17 (0.59,2.31)	1.59 (0.55,4.64)
X-ray recorded	116 (18.4)	24 (7.3)	2.82 (1.21,6.56)	0.43 (0.06,3.01)	0.49 (0.01,16.2)
Core interventions					
OA information provision	549 (87.1)	267 (81.4)	2.21 (1.03,4.74)	1.34 (0.57,3.12)	1.34 (0.57,3.12)
Written OA information	292 (46.3)	12 (3.7)	25.3 (9.50,67.6)	**↑ 24.5 (8.29,72.6)**	**↑ 29.5 (8.52,102)**
Exercise advice provision	522 (82.9)	245 (74.7)	1.94 (1.02,3.66)	1.92 (0.90,4.09)	1.92 (0.90,4.09)
Written exercise advice	230 (36.5)	7 (2.1)	27.5 (8.50,89.2)	**↑ 21.8 (6.81,69.6)**	**↑ 23.5 (6.94,79.5)**
Weight loss advice provision[c]	327 (74.5)	130 (64.0)	2.01 (0.95,4.24)	1.02 (0.45,2.30)	1.02 (0.45,2.30)
Written weight loss advice[c]	99 (22.6)	2 (1.0)	20.0 (4.23,94.7)	**↑ 17.6 (3.28,94.0)**	**↑ 17.6 (3.29,94.0)**
Non-pharmacological management					
Consideration of physiotherapy referral	94 (14.9)	65 (19.8)	1.33 (0.71,2.47)	0.77 (0.38,1.59)	0.77 (0.38,1.59)
Physiotherapy referral made[d]	68 (108)	13 (4.0)	5.28 (2.07,13.5)	**↑ 4.87 (1.83,12.9)**	**↑ 4.87 (1.83,12.9)**

(con't)

Outcome measure	Number of patients with outcome *n* (%)		Odds of outcome (intervention vs. control arm) (95% CI)		
	Intervention arm	Control arm	Unadjusted two-level model[f]	Adjusted two-level model[f]	Adjusted three-level model[g]
Pharmacological management					
Consideration of paracetamol use	549 (87.1)	282 (86.0)	2.89 (1.21,6.92)	1.91 (0.77,4.75)	1.91 (0.77,4.75)
Paracetamol prescribed	163 (25.9)	53 (16.2)	1.97 (1.24,3.12)	↑ **1.76 (1.06,2.92)**	↑ **1.76 (1.06,2.92)**
Consideration of topical NSAID use	496 (78.7)	274 (83.5)	1.19 (0.54,2.60)	1.17 (0.48,2.81)	0.95 (0.32,2.83)
Topical NSAID prescribed	211 (33.5)	90 (27.4)	1.27 (0.74,2.20)	1.09 (0.62,1.92)	1.09 (0.62,1.92)
Oral NSAID prescribed	102 (16.2)	55 (16.8)	0.90 (0.54,1.47)	0.65 (0.39,1.09)	0.60 (0.33,1.08)
Oral NSAID prescribed in the presence of a relative comorbid contraindication	30 (11.7)	14 (11.4)	1.05 (0.41,2.64)	0.86 (0.30,2.49)	0.86 (0.30,2.49)
Gastroprotection prescribed (PPI)[e]	46 (45.1)	23 (41.8)	1.08 (0.55,2.15)	1.50 (0.35,6.52)	0.81 (0.11,6.27)
Opioid prescribed	227 (36.0)	97 (29.6)	1.33 (0.91,1.94)	0.88 (0.62,1.25)	0.88 (0.62,1.25)

Adjusted for diagnostic group, sex, age band, site of disease, BMI status, total morbidity, multiple clinical OA consultation s, consultation with a clinician holding more than the median number of index consultations and practice pre-trial achievement except [a]adjusted for diagnostic group, sex, age, practice pre-trial achievement; [b]MQL1 model; [c]in people known to be overweight at the index consultation; [d]PQL1 model; [e]denominator: those prescribed oral NSAIDs; [f]patients within clinicians; [g]patients within clinicians within practices

Odds of outcome in intervention practices compared to control, trial period (95% CI) in patients with at least one recorded template entry