Internal Clinical Guidelines Team

Spondyloarthritis in over 16s

Spondyloarthritis in over 16s: diagnosis and management

Clinical Guideline

Methods, evidence and recommendations

September 2016

Commissioned by the National Institute for Health and Care Excellence

1

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Copyright
National Institute for Health and Care Excellence, 2016

Contents

1

1 Summary section

Spondyloarthritis encompasses a group of inflammatory conditions with some shared features, including extra-articular manifestations. Both peripheral and axial joints can be affected. The spondyloarthritides are distinct from rheumatoid arthritis but are as important to recognise and manage early in their presentation to improve health outcomes.

The majority of people with these conditions have either psoriatic arthritis or axial spondyloarthritis, which includes ankylosing spondylitis. Ankylosing spondylitis and non-radiographic axial spondyloarthritis primarily affect the spine, in particular the sacroiliac joint. Both conditions manifest in similar ways; the primary classification difference is whether sacroiliitis is detectable on X-ray.

Psoriatic arthritis may manifest in a number of different patterns. These include predominant involvement of small joints in the hands and feet, predominant large joint involvement particularly in the knees or combinations of these. Psoriatic arthritis may also involve the axial joints, and inflammation of the entheses and/or finger and toe joints. Skin and nail involvement may not be present at diagnosis and in its absence, a family history of psoriasis is required to meet the diagnostic criteria.

Less common subgroups are enteropathic spondyloarthritis, which is associated with inflammatory bowel disease (Crohn's disease and ulcerative colitis), and reactive arthritis, which can occur in people following gastrointestinal or genitourinary infections.

The final subgroup is people who have undifferentiated spondyloarthritis. These people generally have an asymmetrical oligoarticular (fewer than 5 involved joints) arthritis, often involving the knees. They do not meet the diagnostic criteria of the other subgroups at presentation but their disease may evolve to do so at a later stage.

This guideline also includes people who are 16 years or older with axial or peripheral symptoms who have previously been diagnosed with juvenile idiopathic arthritis.

Healthcare professionals in non-specialist settings do not always recognise the signs and symptoms of spondyloarthritis, particularly spinal symptoms, which may be mistakenly attributed to other causes of low back pain. This can lead to substantial delays in diagnosis and treatment with consequent disease progression and disability. This guideline seeks to raise awareness of the features of spondyloarthritis and provide clear advice on what action to take when people with signs and symptoms first present in healthcare settings.

This guideline also provides advice on the interventions available to people with spondyloarthritis. These include pharmacological and non-pharmacological treatments, and surgery. The guidance also provides advice on how care for people with spondyloarthritis should be organised across healthcare settings, and what information and support should be provided.

1 GDG membership and ICG technical team

1.1 Guideline Development Group 2016

Gary McVeigh (Guideline Chair)
Professor of Cardiovascular Medicine, Queen's University Belfast & Consultant Physician, Belfast Health and Social Care Trust

Alexander D. L. Baker (co-opted expert member)
Consultant Orthopaedic Spine Surgeon, Lancashire Teaching Hospitals NHS Foundation Trust

Issak Bhojani (Until August 2015)
GP, Shifa Surgery, Blackburn

Nicky Bassett-Burr (co-opted expert member)
Advanced Occupational Therapist, Western Sussex NHS Trust

David Chandler
Patient/carer member

Seau Cheung (co-opted expert member)
Dermatologist, The Royal Wolverhampton NHS Trust

Debbie Cook
Patient/carer member

Charlotte Davis
Rheumatology Clinical Nurse Specialist, Leeds Teaching Hospitals NHS Trust

Alastair Denniston (co-opted expert member)
Consultant Ophthalmologist, University Hospitals Birmingham NHS Foundation Trust & Honorary Reader, University of Birmingham

Nicola Goodson
Rheumatologist, Aintree University Hospital NHS Foundation Trust

Tina Hawkins
Advanced Clinical Pharmacist (Rheumatology), Leeds Teaching Hospital NHS Trust

Philip Helliwell (Until October 2014)
Rheumatologist, Bradford Teaching Hospitals NHS Foundation Trust

Amanda Isdale
Consultant Rheumatologist and Honorary Lecturer, York Teaching Hospital NHS Foundation Trust

Carol McCrum
Consultant Physiotherapist, East Sussex Healthcare NHS Trust

Jon Packham
Consultant Rheumatologist, Staffordshire and Stoke on Trent Partnership NHS and University Hospital of North Midlands NHS Trust. Honorary Senior Clinical Lecturer, Institute of Applied Clinical Sciences, Keele University

Claire Strudwicke (Until May 2015)
Patient/carer member

1 **Louise Warburton**
2 Medical Director, Telford Musculoskeletal Service. GP, Stirchley Medical Practice, Telford.
3 Senior Lecturer at Keele University

4 **Philip O'Connor (co-opted expert member)**
5 Consultant Musculoskeletal Radiologist, Leeds Teaching Hospitals NHS Trust

6 **Tim Orchard (co-opted expert member)**
7 Consultant Physician & Gastroenterologist, Imperial College Healthcare NHS Trust &
8 Professor of Gastroenterology, Imperial College London

9 **Winston Rennie (co-opted expert member)**
10 Consultant Musculoskeletal Radiologist, Leicester Royal Infirmary

11 **Debajit Sen (co-opted expert member)**
12 Consultant in Paediatric, Adolescent and Adult Rheumatology, University College London
13 and Great Ormond Street NHS Foundation Trusts

14 For a full list of guideline development group and service delivery group declarations of
15 interest, see Appendix A.

1.2 Internal Clinical Guidelines Team

17 **Sara Buckner (February to May 2015)**
18 Technical Analyst

19 **Jemma Dean** (In addition: Lynda Ayiku, Elizabeth Barrett and Wesley Hubbard)
20 Information Specialist

21 **Maggie Derry (Until September 2015)**
22 Project Manager

23 **Sue Ellerby**
24 Consultant Clinical Advisor

25 **Nicole Elliot (Until June 2014)**
26 Associate Director

27 **Lucy Hoppe (June to August 2015)**
28 Technical Analyst

29 **Rachel Houten (Until December 2014 and from September 2015 to July 2016)**
30 Health Economist

31 **Holly Irwin (September 2015 to November 2015)**
32 Project Manager

33 **Neel Jain (July to August 2015)**
34 Medical Student

35 **Katherine McAllister (from September 2014)**
36 Technical Analyst

37 **Hugh McGuire (Until December 2015)**
38 Technical Adviser

39 **Vonda Murray (From November 2015)**
40 Project Manager

41 **Shelly Patel**
42 Medicines Adviser

1   **Joshua Pink (From February 2016)**
2   Technical Adviser

3   **Robby Richey (August 2015 to January 2016)**
4   Technical Analyst

5   **Gabriel Rogers**
6   Technical Adviser, Health Economics

7   **Sue Spiers (from June 2014)**
8   Associate Director

9   **Sharlene Ting (January to March 2015)**
10   Technical Analyst

11   **Steven Ward (from December 2014 to September 2015)**
12   Health Economist

13   1.3   Peer review

14   Dr Issak Bhojani

2 Strength of recommendations

Some recommendations can be made with more certainty than others. The Guideline Committee makes a recommendation based on the trade-off between the benefits and harms of an intervention, taking into account the quality of the underpinning evidence. For some interventions, the Guideline Committee is confident that, given the information it has looked at, most patients would choose the intervention. The wording used in the recommendations in this guideline denotes the certainty with which the recommendation is made (the strength of the recommendation).

For all recommendations, NICE expects that there is discussion with the patient about the risks and benefits of the interventions, and their values and preferences. This discussion aims to help them to reach a fully informed decision (see also 'Patient-centred care').

Interventions that must (or must not) be used

We usually use 'must' or 'must not' only if there is a legal duty to apply the recommendation. Occasionally we use 'must' (or 'must not') if the consequences of not following the recommendation could be extremely serious or potentially life threatening.

Interventions that should (or should not) be used – a 'strong' recommendation

We use 'offer' (and similar words such as 'refer' or 'advise') when we are confident that, for the vast majority of patients, an intervention will do more good than harm, and be cost effective. We use similar forms of words (for example, 'Do not offer…') when we are confident that an intervention will not be of benefit for most patients.

Interventions that could be used

We use 'consider' when we are confident that an intervention will do more good than harm for most patients, and be cost effective, but other options may be similarly cost effective. The choice of intervention, and whether or not to have the intervention at all, is more likely to depend on the patient's values and preferences than for a strong recommendation, and so the healthcare professional should spend more time considering and discussing the options with the patient.

3 Methods

This guideline was developed in accordance with the process set out in 'The guidelines manual (2012)'. There is more information about how NICE clinical guidelines are developed on the NICE website. A booklet, 'How NICE clinical guidelines are developed: an overview for stakeholders, the public and the NHS' is available. In instances where the guidelines manual does not provide advice, additional methods are used and are described below.

3.1 Evidence synthesis and meta-analyses

Where possible, meta-analyses were conducted to combine the results of studies for each outcome. For continuous outcomes, where changes from baseline data were reported in the trials and were accompanied by a measure of spread (for example standard deviation), these were extracted and used in the meta-analysis. Where measures of spread for change from baseline values were not reported, the corresponding values at study end were used and were combined with change from baseline values to produce summary estimates of effect. These studies were assessed to ensure that baseline values were balanced across the treatment groups; if there were significant differences at baseline these studies were not included in any meta-analysis and were reported separately.

3.2 Evidence of effectiveness of interventions

3.2.1 Quality assessment

GRADE was used to assess the quality of evidence for the selected outcomes as specified in 'The guidelines manual (2012)'.Where RCTs are possible, these are initially rated as high quality and the quality of the evidence for each outcome was downgraded or not from this initial point. If non-RCT evidence was included for intervention-type systematic reviews then these are initially rated as low quality and the quality of the evidence for each outcome was further downgraded or not from this point.

3.2.2 Methods for combining intervention evidence

Meta-analyses of interventional data were conducted with reference to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al. 2011).

A pooled relative risk was calculated for dichotomous outcomes (using the Mantel–Haenszel method).

Random-effects models (der Simonian and Laird) were fitted for all syntheses, as a conservative approach that reflected the underlying clinical heterogeneity of interventions (for example, complex non-pharmacological programmes), regardless of whether such heterogeneity could be statistically identified.

Meta-analyses were performed in Cochrane Review Manager v5.3.

3.2.3 GRADE for pairwise meta-analyses for interventional evidence

The quality of the evidence for each outcome was downgraded where appropriate for the reasons outlined in Table 1

Table 1: Rationale for downgrading quality of evidence for intervention studies

GRADE criteria	Example reasons for downgrading quality
Risk of bias	Concerns about the design or execution of the study, including concealment of allocation, blinding, loss to follow up. These were identified using intervention

GRADE criteria	Example reasons for downgrading quality
	checklists in the NICE guidelines manual (2012).
Inconsistency	Concerns about inconsistency of effects across studies, occurring when there is variability in the treatment effect demonstrated across studies (heterogeneity). This was assessed using the statistic, I^2 where $I^2 < 33\%$ was categorised as no inconsistency, I^2 between 33% and 66% was categorised as serious inconsistency, and $I^2 \geq 67\%$ was categorised as very serious inconsistency
Indirectness	Concerns about the population, intervention and outcome in the included studies and how directly these variables could address the specific review question.
Imprecision	Uncertainty around the estimate of effect, for example when the confidence intervals are wide and cross the lines of clinically significant effect. This reflects the confidence in the estimate of effect. If no minimum clinically important difference could be defined, outcomes were downgraded if the confidence interval crossed the line of no effect (i.e. the outcome was not statistically significant).

3.3 Diagnostic evidence

A number of questions in this guideline relied on diagnostic accuracy evidence. It should be noted that the term 'diagnostic accuracy' does not necessarily imply that the data – and the features they represent – should be used for strictly diagnostic purposes; indeed, these data span questions regarding suspicion, referral and formal diagnosis, in this guideline. From a methodological point of view, diagnostic accuracy data may be classified as any data in which a feature – be it a symptom, a risk factor, a test result or the output of some algorithm that combines many such features – is observed in some people who have the condition of interest and some people who do not. Such data either explicitly provide, or can be manipulated to generate, a 2x2 classification of true positives and false negatives (in people who, according to the reference standard, truly have spondyloarthritis) and false positives and true negatives (in people who, according to the reference standard, do not).

The 'raw' 2x2 data can be summarised in a variety of ways. Those that were used for decision making in this guideline are as follows:

- **Positive likelihood ratios** describe how many times more likely positive features are in people with spondyloarthritis compared with people without spondyloarthritis. Values greater than 1 indicate that a positive result makes spondyloarthritis more likely.
  - $LR^+ = (TP/[TP+FN])/(FP/[FP+TN])$
- **Negative likelihood ratios** describe how many times less likely negative features are in people with spondyloarthritis compared with people without spondyloarthritis. Values less than 1 indicate that a negative result makes spondyloarthritis less likely.
  - $LR^- = (FN/[TP+FN])/(TN/[FP+TN])$
- **Sensitivity** is the probability that the feature will be positive in a person with spondyloarthritis.
  - sensitivity = TP/(TP+FN)
- **Specificity** is the probability that the feature will be negative in a person without spondyloarthritis.
  - sensitivity = TN/(FP+TN)

The GDG put particular priority on positive and negative likelihood ratios in their decision making. The following schema, adapted from the suggestions of Jaeschke et al. (1994), was used to interpret findings.

Table 2: Interpretation of likelihood ratios

Value of likelihood ratio	Interpretation
LR ≤ 0.1	**Very large** decrease in probability of disease
0.1 < LR ≤ 0.2	**Large** decrease in probability of disease
0.2 < LR ≤ 0.5	**Moderate** decrease in probability of disease
0.5 < LR ≤ 1.0	**Slight** decrease in probability of disease
1.0 < LR < 2.0	**Slight** increase in probability of disease
2.0 ≤ LR < 5.0	**Moderate** increase in probability of disease
5.0 ≤ LR < 10.0	**Large** increase in probability of disease
LR ≥ 10.0	**Very large** increase in probability of disease

This schema has the effect of setting a minimally important difference for a positive likelihood ratio at 2, and a corresponding minimally important difference for negative likelihood ratios at 0.5. Likelihood ratios (whether positive or negative) falling between these thresholds were judged to indicate no meaningful change to probability of disease.

3.3.1   Methods for combining diagnostic evidence

Meta-analysis of diagnostic test accuracy data was conducted with reference to the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (Deeks et al. 2010).

All diagnostic syntheses were doubly stratified:

- by presenting symptomatology:
  - o people with predominantly axial symptoms
  - o people with predominantly peripheral symptoms
  - o mixed studies including people with axial and/or peripheral symptoms
- and by reference standard:
  - o expert clinician diagnosis
  - o diagnosis according to published criteria

Each data point was categorised according to these features to create up to 6 substrata with separate summary estimates

Separate pooling was performed for positive likelihood ratios, negative likelihood ratios, sensitivity and specificity. This is a theoretically suboptimal approach, as correlations exist between positive and negative likelihood ratios and between sensitivities and specificities, and these are not accounted for in this method. Superior methods of synthesis exist, which incorporate the related outcomes in a bivariate model. However, in order to estimate the parameters of such models, it is necessary to provide a minimum amount of observations (that is, studies); for example, the Stata bivariate meta-analysis command `metandi` will not attempt synthesis unless there are 5 or more studies on which to base outputs. Sufficient data for this approach were very seldom available in the evidence-base for this guideline, as it would have been necessary to meet this minimum criterion for each substratum of the analysis (equivalently, covariates for presenting symptomatology and reference standard could have been introduced to a bivariate model, but this would have increased data requirements to a similar degree).

For these reasons, there was no feasible alternative to independent pooling of each summary statistic, even though this may somewhat underestimate test accuracy (see Deeks 2001).

1 Random-effects models (der Simonian and Laird) were fitted for all syntheses, as
2 recommended in the Cochrane Handbook for Systematic Reviews of Diagnostic Test
3 Accuracy (Deeks et al. 2010).

4 Diagnostic meta-analyses were performed in Microsoft Excel.

3.3.2 Modified GRADE for diagnostic evidence

6 GRADE has not been developed for use with diagnostic studies; therefore a modified
7 approach was applied using the GRADE framework.

8 GRADE assessments were only undertaken for positive and negative likelihood ratios, as
9 these were preferred by the GDG as summary measures of diagnostic accuracy.

10 Cross-sectional and cohort studies were initially rated as high-quality evidence if well
11 conducted, and then downgraded according to the standard GRADE criteria (risk of bias,
12 inconsistency, imprecision and indirectness) as detailed in Table 3 below.

13 **Table 3: Rationale for downgrading quality of evidence for diagnostic questions**

GRADE criteria	Example reasons for downgrading quality
Risk of bias	This includes limitations in the design or execution of the study. Assessment was based on the QUADAS 2 checklist; studies were downgraded if there was evidence of bias in at least 2 domains or of serious bias in 1. Particular attention was paid to non-consecutive recruitment of participants and blinding of reference standard (in retrospective studies where the final diagnosis was known).
	Datasets with more than 1 study were downgraded for risk of bias if one-third or more of the weight in meta-analysis came from studies that had been judged to be at serious risk of bias (that is, datasets that did not have at least twice as much evidence from studies at low risk of bias as from studies at serious risk of bias were downgraded).
Inconsistency	The quality of the evidence was downgraded if there were concerns about Inconsistency of effects across studies. The I^2 statistic was used – values <50% were categorised as subject to no serious inconsistency and values ≥50% were categorised as suffering from serious inconsistency.
	This approach is somewhat less conservative than that used in intervention studies (see Table 1), for the reason that heterogeneity is an unavoidable feature of diagnostic syntheses. The Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy states that 'Heterogeneity is to be expected in meta-analyses of diagnostic test accuracy. A consequence of this is that meta-analyses of [diagnostic] accuracy studies tend to focus on computing *average* rather than *common* effects…. In [diagnostic] accuracy reviews large differences are commonly noted between studies, too big to be explained by chance, indicating that actual test accuracy varies between the included studies, or that there is heterogeneity in test accuracy' (Deeks et al. 2010). For these reasons, it was considered unnecessarily conservative to doubly downgrade analyses with I^2 values ≥67%, or downgrade analyses with I^2 values of 33–50% at all.
Indirectness	The quality of the evidence was downgraded if there were concerns about the population, index feature or reference standard in the included studies and how directly these variables could address the specific review question. Studies were automatically downgraded if they had a reference standard of published criteria, as this was recognised by the GDG as inferior to their preferred standard of expert clinician diagnosis.
	Datasets with more than 1 study were downgraded for risk of bias if one-third or more of the weight in meta-analysis came from studies that had been judged to suffer from serious indirectness (that is, datasets that did not have at least twice as much evidence from directly relevant studies as from studies with serious indirectness were downgraded).

GRADE criteria	Example reasons for downgrading quality
Imprecision	The quality of the evidence was downgraded if there is important uncertainty around the estimate of effect, for example when the confidence intervals are cross the lines of clinically significant effect. In line with the definitions in Table 2, if the 95% confidence interval for a positive likelihood ratio spanned 2, the outcome was downgraded, as the data were deemed to be consistent with a meaningful increase in risk and no meaningful predictive value. Similarly, negative likelihood ratios that spanned 0.5 led to downgrading for serious imprecision. Any likelihood ratios that spanned **both** 0.5 and 2 were downgraded twice, as suffering from very serious imprecision.

3.4 Methods for combining direct and indirect evidence (network meta-analysis)

Conventional 'pairwise' meta-analysis involves the statistical combination of direct evidence about pairs of interventions that originate from two or more separate studies (for example, where there are two or more studies comparing A vs B).

In situations where there are more than two interventions, pairwise meta-analysis of the direct evidence alone is of limited use. This is because multiple pairwise comparisons need to be performed to analyse each pair of interventions in the evidence, and these results can be difficult to interpret. Furthermore, direct evidence about interventions of interest may not be available. For example studies may compare A vs B and B vs C, but there may be no direct evidence comparing A vs C. Network meta-analysis overcomes these problems by combining all evidence into a single, internally consistent model, synthesising data from direct and indirect comparisons, and providing estimates of relative effectiveness for all comparators and the ranking of different interventions.

The evidence in section 5.1 of this guideline was analysed using network meta-analysis, to inform decisions about pharmacological management of axial spondyloarthritis.

3.4.1 Synthesis

Hierarchical Bayesian Network Meta-Analysis (NMA) was performed using WinBUGS version 1.4.3. The models used reflected the recommendations of the NICE Decision Support Unit's Technical Support Documents (TSDs) on evidence synthesis, particularly TSD 2 ('A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials'; see http://www.nicedsu.org.uk). The WinBUGS code provided in the appendices of TSD 2 was used without substantive alteration to specify synthesis models.

Results were reported summarising 70,000 samples from the posterior distribution of each model, having first run and discarded 35,000 'burn-in' iterations. Three separate chains with different initial values were used.

3.4.2 Prior distributions

Non-informative prior distributions were used in all models. Unless otherwise specified, trial-specific baselines and treatment effects were assigned N(0, 1000) priors, and the between-trial standard deviations used in random-effects models were given U(0, 5) priors. These are consistent with the recommendations in TSD 2 for dichotomous outcomes.

3.4.3 Applying GRADE to network meta-analysis

The use of GRADE to assess the quality of studies addressing a particular review question for pairwise comparisons of interventions is relatively established. However, the use of

GRADE to assess the quality of evidence across a network meta-analysis is still a developing methodology. While most criteria for pairwise meta-analyses still apply, it is important to adapt some of the criteria to take into consideration additional factors, such as how each 'link' or pairwise comparison within the network applies to the others. As a result, the following was used when modifying the GRADE framework to a network meta-analysis. It is designed to provide a single overall quality rating for an NMA, which can then be combined with pairwise quality ratings for individual comparisons (if appropriate), to judge the overall strength of evidence for each comparison.

Risk of bias

For network meta-analyses with a large proportion of studies that were judged to be susceptible to bias, the following downgrading decision rule was applied.

- If 50% or more studies in the network were inadequate or unclear for a particular parameter of quality, the outcome was downgraded by 1 level.

Inconsistency

Decisions between fixed and random effects model specifications were made based on the Deviance Information Criterion (DIC). The network was downgraded for inconsistency if the DIC suggested a random effects model should be preferred.

Indirectness

As with pairwise meta-analyses, studies included in a network were assessed for how well they fit the PICO (population, intervention, comparator, outcome) specified in the review protocol.

Imprecision

Imprecision was assessed for a number of variables:

- Sufficient head-to-head trials in the network.
- Sufficient number of studies to form the network (if there was a high proportion of 'links' formed with only 1 trial, the outcome was downgraded).
- Overall certainty/uncertainty of the effect estimates (size of credible intervals, including for each drug compared with the reference option).

The overall outcome was downgraded for imprecision if the above factors meant it was not possible to differentiate between any meaningfully distinct options in the network (e.g. treatment versus no treatment).

4 Recognition, referral and diagnosis

Spondyloarthritis encompasses a number of related conditions with different manifestations. This makes the task of developing a comprehensive, clinically useful guideline for diagnosing and managing spondyloarthritis more difficult, particularly because there is very little evidence available. The development of rheumatoid arthritis guidelines has improved awareness of early inflammatory arthritis resembling possible rheumatoid arthritis, and prompted the development of rapid access pathways between primary and secondary care. In contrast, the spondyloarthritides have not attracted similar attention, and to date there have been no guidelines to cover this group of disorders.

Spondyloarthritis forms a significant proportion of inflammatory musculoskeletal conditions. Delays in correctly identifying and diagnosing spondyloarthritis can result in significant morbidity and reduced ability to work, and wastes valuable resources on inappropriate investigations and treatments. This is particularly true in axial spondyloarthritis which can present with insidious symptoms that may be difficult to differentiate from simple mechanical back pain, unless appropriate assessment and imaging is undertaken. For example, it can take 8 years to develop sacroiliitis detectable with plain film radiography. Using this sign as a criterion for diagnosis will significantly delay early diagnosis and potentially miss the treatment window for preventing irreparable damage.

The first challenge is to raise healthcare professionals' awareness and understanding about spondyloarthritis, to help them recognise risk factors and early symptoms and signs. This can be challenging because spondyloarthritis includes a number of heterogeneous conditions that affect both peripheral and axial joints, often with extra-articular features. Examples include psoriatic arthritis, reactive arthritis and conditions primarily affecting the spine, with ankylosing spondylitis as the most severe form. GPs or other healthcare professionals may only see one patient each year with the presenting features of spondyloarthritis. This means that in primary care the necessary skill levels are difficult to acquire and maintain.

A review of the evidence for when to suspect early spondyloarthritis (which would then prompt specialist referral to confirm the diagnosis) is presented. Raising awareness amongst clinicians of the early signs and symptoms of spondyloarthritis should improve the likelihood of a person with spondyloarthritis receiving a prompt and correct diagnosis and appropriate management.

A further issue is the variation in referral strategies across the UK. Some areas use referral to interface musculoskeletal services as a triage process before people with suspected spondyloarthritis access rheumatology services. Within these interface services, many healthcare professionals have the necessary skills to diagnose spondyloarthritis, but not all do. In addition, there is no national accreditation of such services or agreed core competencies for the healthcare professionals working within these services. Therefore, there may be a further inadvertent delay in onward referral to specialist rheumatology services. Ascertaining the most appropriate and cost-effective care pathway is also necessary, and the available evidence is reviewed.

Once spondyloarthritis is suspected, there are a range of tests and tools which may be used to investigate further in order to reach a diagnosis. No single sign, symptom or test result has proved useful for diagnosis in isolation of other information. Eliciting information from the person about their symptoms and risk factors (including family history of spondyloarthritis and associated extra-articular conditions) is the starting point. Investigations which may be considered include imaging of affected joints, and testing for circulating and genetic biomarkers. The gene HLA-B27 is well known to be associated with spondyloarthritis, particularly axial disease, though this knowledge has led to the misconception that testing positive for this marker is an essential requirement for a positive diagnosis. Similarly, where sacroiliitis on X-ray used to be a mandatory sign for axial spondyloarthritis to be diagnosed and labelled as ankylosing spondyloarthritis, the use of MRI has enabled greater detection of

1 sacroiliitis and inflammatory back pain, giving rise to the diagnosis of 'non-radiographic axial
2 spondyloarthritis'. A range of diagnostic tools and models exists to support clinical decision
3 making, but evaluation of these shows that different tools may yield different results across a
4 population.

5 In conclusion, there are many obstacles to the prompt diagnosis of spondyloarthritis.
6 Although there is a dearth of quality evidence for best practice, specialist spondyloarthritis
7 services are well-established within many rheumatology/musculoskeletal departments.
8 Patient support groups, particularly the National Ankylosing Spondylitis Society (NASS) and
9 the Psoriasis and Psoriatic Arthritis Alliance (PAPAA) also provide an important portal for
10 people to obtain information and guidance on local services available and managing these
11 conditions.
12

4.1 Identifying new cases of spondyloarthritis

Review questions 1, 2, 12, 6 and 3

- 1. What signs and symptoms should prompt a healthcare professional to think of spondyloarthritis?
- 2. What risk factors should increase suspicion of spondyloarthritis?
- 12. What are the indications (signs, risk factors, test or scan findings) for referral for specialist advice at initial diagnosis?
- 6. What is the comparative effectiveness of different referral strategies in diagnosing spondyloarthritis?
- 3. What are the obstacles to a prompt diagnosis of spondyloarthritis?

4.1.1 Evidence review

The aim of these review questions was to improve the recognition of spondyloarthritis at the initial point of contact with health care professionals, and thereby increase the proportion of people who correctly receive a referral for diagnosis in a specialist setting.

Table 4: PICO table: signs and symptoms of spondyloarthritis

Population	People (aged 16 years and over) with suspected spondyloarthritis, or people with diagnosed spondyloarthritis whose presenting symptoms are being studied
Intervention	Signs and symptoms including: **Axial:** low/general back pain (>3 months), onset of back pain age<45, spinal fusion, neck pain, morning stiffness, stiffness, limited mobility inflammatory bowel disease, psoriasis, uveitis, site-specific inflammation/pain, enthesitis, fatigue, signs on imaging, response to NSAIDs, buttock pain **Peripheral:** joint pain and swelling, oligoarthritis, enthesitis, dactylitis, inflammatory bowel disease, psoriasis, uveitis, examination showing suspected persistent synovitis of undetermined cause, site-specific inflammation/pain, nail involvement, fatigue, morning stiffness, signs on imaging. **Reactive arthritis:** urethritis, keratoderma blennorrhagica, conjunctivitis, balanitis, soft palate ulceration
Comparator	Expert clinician diagnosis of spondyloarthritis was considered the preferred reference standard, with diagnosis using any specified criteria as the next preference.
Outcomes	• Sensitivity • Specificity • Positive likelihood ratio • Negative likelihood ratio • Positive predictive value • Negative predictive value • Diagnostic odds ratio

Table 5: PICO table: risk factors for spondyloarthritis

Population	People (aged 16 years and over) with suspected spondyloarthritis, or people with diagnosed spondyloarthritis whose presenting symptoms are being studied
Intervention	family history, HLA-B27 positive, history of psoriasis, history of IBD, history of uveitis, history of ReA, history of JIA (enthesitis/psoriatic), recent enteric or genitourinary infection, onset under age 45 (axial)
Comparator	Expert clinician diagnosis of spondyloarthritis was considered the preferred reference standard, with diagnosis using any specified criteria as the next preference.

Outcomes	• Sensitivity • Specificity • Positive likelihood ratio • Negative likelihood ratio • Positive predictive value • Negative predictive value • Diagnostic odds ratio

1

Table 6: PICO table: indications for referral

Population	People (aged 16 years and over) with suspected spondyloarthritis
Intervention	Indications to include: • Inflammatory lower back pain (axial) of at least 3 months duration often with insidious onset • Joint/tendon pain (axial or peripheral)/swelling (peripheral) • Morning stiffness or stiffness improving with exercise • Elevated ESR/CRP • HLA-B27 positive • Family history • Presence of extra-articular symptoms (uveitis, psoriasis, IBD) • Radiographic/imaging signs if available • NSAID responsiveness • Reactive arthritis
Comparator	Expert clinician diagnosis of spondyloarthritis was considered the preferred reference standard, with diagnosis using any specified criteria as the next preference.
Outcomes	• Sensitivity • Specificity • Positive likelihood ratio • Negative likelihood ratio • Positive predictive value • Negative predictive value • Diagnostic odds ratio

2

Table 7: PICO table: comparative effectiveness of referral criteria

Population	People suspected of having spondyloarthritis or people with inflammatory back pain symptoms
Intervention	Referral strategy/ protocol/proforma/pathway
Comparator	Any other referral strategy
Outcomes	• percentage of referrals correctly diagnosed as spondyloarthritis • time taken from symptoms to diagnosis (not time from referral) • resource use and costs • health-related quality of life • improvement in disease specific outcomes • reduced long term complications and/or skeletal damage

3

Table 8: PICO table: obstacles to diagnosis

Population	Patients with suspected or confirmed diagnosis of spondyloarthritis, healthcare professionals

Interventions	Barriers such as:
	• Lack of patient awareness leading to delayed diagnosis
	• Patients deterred by lack of diagnosis at earlier consultation
	• Lack of health-care professional awareness of chronic inflammatory conditions
	• Lack of health-care professional awareness of complications/co-morbid manifestations of pre-existing inflammatory conditions
	• High consultation rate of lower back pain (mostly mechanical)
	• Lack of cross referrals in secondary care between relevant specialities
	• Over-specialism within rheumatology leading to consultations where relevant comorbidities are not assessed.
	• Lack of multidisciplinary team assessment
	• Lack of access from GPs to (i) HLA-B27 testing (ii) appropriate MRI equipment or protocol
	• Patient gender (under-diagnosis in women)
	• Lack of a biological marker in spondyloarthritis
Comparators	Prompt diagnosis of spondyloarthritis
Outcomes	Time to appointment, number of contacts with health care professionals, health related quality of life, resource use and costs, patient satisfaction, disease burden reduced from both spondyloarthritis and associated conditions, service delivery/organisation

1   For full details of the review protocols please see Appendix C.

2   Cross sectional studies and cohort studies were considered to be the highest quality
3   evidence available to answer questions 1, 2 and 12. Randomised controlled trials were
4   considered the highest quality evidence for question 6. Qualitative studies were the preferred
5   design for question 3.

6   Studies are graded as high in a modified GRADE framework if they are of the preferred study
7   design for that question and are conducted and reported well.

8   Systematic searches for questions 1, 2 and 12 identified 12,797 references, which were
9   screened on their titles and abstracts. Between papers identified from this search and papers
10   identified from other diagnostic searches in this guideline which contained relevant data, a
11   total of 133 studies were retrieved for full text review. 102 of these studies were excluded as
12   they did not meet the eligibility criteria such as inappropriate study design (e.g. case reports,
13   case series with fewer than ten cases) or non-primary studies (e.g. systematic review,
14   editorial).

15   Systematic searches for question 6 identified 1,234 references, which were screened on their
16   titles and abstracts. Between papers identified from this search and papers identified from
17   other diagnostic searches in this guideline which contained relevant data, a total of 5 studies
18   were retrieved for full text review. Three of these studies were excluded as they did not meet
19   the eligibility criteria such as inappropriate study design (e.g. case reports, case series with
20   fewer than ten cases) or non-primary studies (e.g. systematic review, editorial).

21   Systematic searches for question 3 identified 11,413 references, which were screened on
22   their titles and abstracts. Between papers identified from this search and papers identified
23   from other diagnostic searches in this guideline which contained relevant data, a total of 25
24   studies were retrieved for full text review. 20 of these studies were excluded as they did not
25   meet the eligibility criteria such as inappropriate study design (e.g. case reports, case series
26   with fewer than ten cases) or non-primary studies (e.g. systematic review, editorial). An
27   update search carried out near the end of guideline development identified 2 further studies.

28

29

1 Detailed lists of excluded studies and reasons for their exclusion for all questions are
2 provided in Appendix F.

4.1.1.1 Description of included studies

4 Evidence tables for included studies can be found in Appendix E, with GRADE profiles
5 reported in Appendix G.

4.1.1.1.1 *Signs, symptoms and risk factors*

7 Data were identified from cross-sectional studies for the diagnostic value of the following
8 features:

9 • Inflammatory back pain (4 studies ASAS criteria, 2 studies Berlin Criteria, 5 studies Calin
10   criteria, 6 studies with other criteria, 3 studies on back pain in people with other presenting
11   complaints)– axial, peripheral and mixed populations

12 • Age of back pain onset (4 studies) – axial and mixed populations

13 • Morning stiffness (2 studies) – axial and mixed populations

14 • Neck pain (1 study) – axial population

15 • Response to NSAIDs (9 studies) – axial and mixed populations

16 • Enthesitis (14 studies general enthesitis, 5 studies heel enthesitis) – axial, peripheral and
17   mixed populations

18 • Psoriasis (9 studies) – axial and mixed populations

19 • Uveitis/history of uveitis (12 studies) – axial, peripheral and mixed populations

20 • Inflammatory bowel disease (7 studies) – axial and mixed populations

21 • Dactylitis (8 studies) – axial, peripheral and mixed populations

22 • Arthritis (11 studies) – axial, peripheral and mixed populations

23 • Nail disease (5 studies) – peripheral population

24 • Fatigue (2 studies)– peripheral population

25 • Family history of spondyloarthritis (12 studies) – axial, peripheral and mixed populations

26 • Family history of psoriasis (2 studies) – peripheral population

27 • Preceding infection (7 studies) – axial, peripheral and mixed populations

4.1.1.1.2 *Indicators for referral*

29 Four studies (2 studies each based on data from 2 cohort studies) reported the diagnostic
30 accuracy of a range of referral strategies, based on setting differing cut-offs for the number of
31 features on a list people needed to satisfy to meet the referral criteria. Participants were
32 followed up to a definitive diagnosis regardless of whether they met a particular referral
33 criteria.

4.1.1.1.3 *Comparison of referral strategies*

35 Two randomised controlled trials were identified (Poddubnyy 2011 and Sieper 2013) which
36 reported the proportions of people with different final diagnoses (axial spondyloarthritis,
37 possible axial spondyloarthritis and no axial spondyloarthritis) for both a simple and more
38 complex referral strategy.

4.1.1.1.4 *Obstacles to diagnosis*

Reference, diagnosis	Study type, number of participants, country	Study design	Outcomes
Aggarwal (2009)	Cross-sectional	Investigator administered	Delay in diagnosis by clinical

Reference, diagnosis	Study type, number of participants, country	Study design	Outcomes
Ankylosing spondylitis	survey N=70 India	questionnaire with patients at a rheumatology clinic	characteristics, mean diagnosis delay Incorrect diagnoses
Dincer (2008) Ankylosing spondylitis	Mixed methods study N=111 Turkey	Face-to-face interview using questionnaire with patients	Clinical features, mean diagnosis delay
Hajialilo (2014) Ankylosing spondylitis	Cross-sectional survey N=60 Iran	Questions about aspects of their condition to patients from rheumatology clinics	Clinical features, mean diagnosis delay
Jois (2008) Ankylosing spondylitis	Cross-sectional survey N=196 UK	Postal survey with GPs	Recognition of symptoms of inflammatory back pain and associated SpA symptoms
Martindale (2014) Ankylosing spondylitis, axial spondyloarthritis	Mixed methods study N=10 UK	Questionnaires and interviews with participants in a larger cohort study, in 2 rheumatology departments	Themes reported, related to process of their diagnosis
Seo (2015) Axial spondyloarthritis	Mixed methods study N=105 Republic of Korea	Face-to-face interviews, review of medical records, with patients in a rheumatology clinic	Alternative diagnosis, factors related to diagnosis >8years
Slodobin (2011) Axial spondyloarthritis	Cross-sectional survey N=151 Israel	Data collected during recruitment visit or from chart review	Features of axial spondyloarthritis by gender Clinical features, diagnosis delay in categories
van Onna (2014) Axial spondyloarthritis	Interviews N=10 The Netherlands	Semi-structured interviews with GPs	Themes reported, knowledge, beliefs and experiences about inflammatory back pain and axial spondyloarthritis

4.1.1.2 Variations from protocol

A specific search was conducted for each of these questions or groups of questions, to identify studies which provided data on recognition and referral of suspected spondyloarthritis. However, in the course of conducting other diagnostic utility questions for this guideline, further data was identified which was contained incidentally in other studies. Across all of the searches conducted for any of the diagnostic questions, any relevant data was extracted if the study met the eligibility criteria, regardless of whether it was identified in the specific search for this study.

4.1.1.3 Minimal clinically important differences

Minimal clinically important differences were considered in two contexts when interpreting the diagnostic evidence in this guideline. When considering individual factors in isolation, it was agreed in collaboration with the GDG that a positive likelihood ratio of 2 would constitute significant diagnostic value. Therefore, when interpreting the diagnostic accuracy results for single factors in isolation, something would only be considered to have diagnostic value if the

1 result was statistically significant at the 95% confidence level, and the point estimate of the
2 positive likelihood ratio was greater than 2 (or equivalently, value at ruling out the disease if
3 the negative likelihood ratio was less than 0.5).

4 When considering models containing multiple factors, the individual predictive value of each
5 factor in isolation was considered not be a meaningful measure, as if the joint effect of a
6 number of factors, which may have only limited diagnostic value individually, is to create an
7 overall algorithm which is highly predictive, considering the diagnostic utility of the individual
8 factors is no longer relevant. Further, there may well be correlations/interactions between
9 factors which means the overall diagnostic value of two factors may be considerably different
10 (in either direction) from the value one would predict it to be simply assuming independence
11 of the individual elements. Therefore, MCIDs were not considered as part of the process of
12 assessing algorithms containing multiple factors.

4.1.2 Health economic evidence

13

4.1.2.1 Systematic review of published literature

14

15 A literature search was conducted jointly for all review questions in this guideline by applying
16 standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
17 total of 9,970 references was retrieved, of which none were retained for these review
18 questions.

4.1.2.2 Original health economic analysis

19

4.1.2.2.1 Methods

20

21 The GDG identified the recognition and appropriate referral of axial spondyloarthritis as its
22 key priority for original health economic analysis. The group advised that delayed diagnosis
23 is a significant issue in all spondyloarthritis, but that people with axial symptoms are subject
24 to particularly damaging delays, invariably because their symptoms are misidentified as
25 mechanical back pain. The GDG emphasised that, if people with axial disease could be
26 identified more reliably when they first present, they would gain access to effective
27 treatments, improving their quality of life and their chances of long-term disease modification.

28 Accordingly, the original model was devised to estimate quality of life and costs (over a
29 lifetime) of people who are and are not correctly referred, having presented with symptoms
30 that might indicate axial spondyloarthritis. It has a 3-month cycle length and a lifetime time
31 horizon, and adopts a patient perspective for outcomes and an NHS perspective for costs, in
32 line with the Guidelines Manual (2012).

33 In reflection of the diagnostic accuracy evidence, the simulated population comprises people
34 with chronic back pain of at least 3 months' duration that began at age 45 or younger. Using
35 data from a large inception study (Rudwaleit et al., 2009), the ankylosing spondylitis (AS)
36 cohort was assumed to be 64% male with an average age of 30.4 (95%CI: 29.0 to 31.8), and
37 the non-radiographic axial spondyloarthritis (nrAxSpA) cohort was 43% male and had a
38 mean age of 33.2 (95%CI: 31.8 to 34.6).

39 Figure 1 provides a schematic depiction of the model structure. It shows that each
40 recognition strategy is modelled in terms of its ability to categorise people into true-positive
41 and true-negative diagnoses (with complementary probabilities of false-negative and false-
42 positive diagnoses, respectively). The long-term costs and QALYs associated with people
43 who do not have spondyloarthritis are not modelled: it is assumed that the specialists to
44 whom false-positive cases are incorrectly referred will identify their true-negative status, so
45 only the costs of specialist diagnostic work-up are modelled. Where true-negative cases are
46 concerned, the choice of referral strategy makes no difference to the future costs and quality

1 of life of people who are correctly identified as not having spondyloarthritis, so there is no
2 need to estimate these.

aTNF=anti-TNF therapy; BSC=best supportive care; FN=false negative; FP=false positive; SpA=spondyloarthritis; TN=true negative; TP=true positive
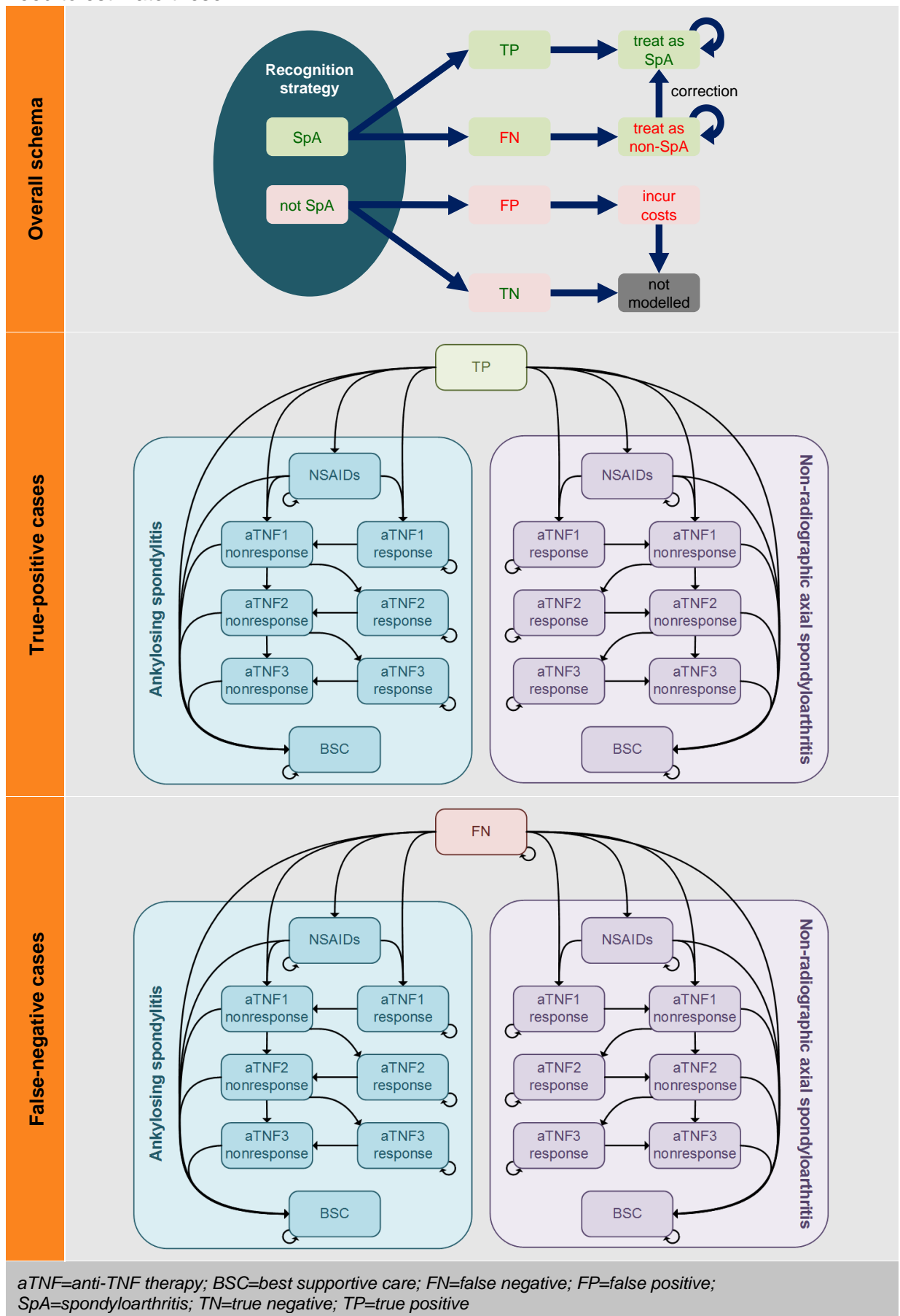
1 **Figure 1: Structure of original cost–utility model**

2 A simplified treatment pathway is assumed for true-positive cases: for most people, first-line
3 treatment is with non-steroidal anti-inflammatory drugs (NSAIDs) (although a proportion of
4 people will be contraindicated and proceed directly to biological disease-modifying anti-
5 rheumatic drugs (DMARDs), unless they are also contraindicated for these, in which case
6 they can only receive best supportive care [BSC]). Up to 3 lines of anti-TNF therapy are
7 modelled, in reflection of technology appraisal guidance TA383. The BSC state is designed
8 to represent the care of people who cannot take – or whose disease no longer responds to –
9 any disease-modifying therapy. A proportion of people within this state are assumed to be
10 referred to a chronic pain management service.

11 The false-negative pathway is identical to the true-positive version, with the critical exception
12 that people remain in the false-negative state (where they are treated as if they have
13 mechanical low-back pain) until their true diagnosis is uncovered. The likelihood of late
14 diagnosis is parameterised used evidence from a survey of 1,630 people with ankylosing
15 spondylitis in the UK.

16 Evidence shows that longer duration of symptoms is associated with a lower chance of
17 response to biological DMARDs (Rudwaleit et al. 2004). Because, when compared with true-
18 positive referrals, the simulated patients who enter the treatment pathway via a diagnostic
19 delay (that is, time spent in the false-negative state) have a longer disease history at the time
20 they start biological DMARDs, they have a lower probability of response.

21 Cases of spondyloarthritis are proportionally allocated between diagnoses of ankylosing
22 spondylitis (AS) and non-radiographic axial spondyloarthritis (nrAxSpA); wherever evidence
23 exists for differential effects between these categories, this is reflected in the model. There
24 are no transitions between the AS and nrAxSpA subgroups. Although it is acknowledged that
25 some people are first diagnosed with non-radiographic disease that subsequently becomes
26 radiographically overt, this is also true of participants in the studies used to populate the
27 nrAxSpA pathway (most notably, RCTs of the effectiveness of biological DMARDs).
28 Therefore, the 'non-radiographic' states in the model can be interpreted as 'axial
29 spondyloarthritis that did not meet radiographic criteria at the time of initial diagnosis'.

30 The average Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) and Bath
31 Ankylosing Spondylitis Functional Index (BASFI) of simulated patients in each state is
32 projected using evidence on natural history and treatment effect. These are used to project
33 quality of life (using a published mapping function [Wailoo et al., 2015]) and background
34 healthcare costs (data from Boonen et al., 2003, as implemented by Corbett et al., 2016).
35 Because BASDAI and BASFI are projected to rise at a steeper trajectory in occult disease
36 than when people are receiving appropriate treatment, people who are diagnosed later have
37 higher values which, in turn, translate into worse quality of life and higher background
38 healthcare costs.

39 Rather than simulating each possible strategy individually, the model calculates the
40 discounted lifetime costs and QALYs expected from true-positive and false-negative cases
41 (as well as the costs associated with false-positive referrals). The costs and effects of any
42 strategy can then be calculated as an average of the relevant outputs, weighted according to
43 the proportion of TP, FN and FP cases the strategy is predicted to produce (which is, in turn,
44 a simple function of the sensitivity and specificity of the strategy and the true prevalence of
45 axial spondyloarthritis in the presenting population).

46 Strategies that were evaluated in the model are described in Table 9. Evidence of the
47 appropriate type – that is, following people with possible AxSpA until final diagnosis,
48 regardless of whether they met particular criteria – is limited, and dominated by reports from
49 2 cohorts (Braun et al. 2011, 2013; van Hoeven et al. 2015).

1 **Table 9: Evaluated strategies**

Study	Strategy	Description
van Hoeven (2015)	>=x	A score of x or more on the CaFaSpA scoring system: • positive ASAS IBP questionnaire (1pt) • family history (1pt) • good response to NSAIDs (1pt) • duration >5yr (0.5pt)
van Hoeven (ASAS) – validation of ASAS referral criteria in CaFaSpA cohort (van Hoeven et al. 2015)	>=x	x or more criteria from the ASAS referral criteria met (as validated in the CaFaSpA cohort): • IBP • arthritis, enthesitis or dactylitis • psoriasis, IBD or uveitis • family history • good response to NSAIDs • elevated CRP or ESR • HLA-B27 positivity • Sacroiliitis on imaging (if available)
van Hoeven (SSB27) – combinations of features assessed in CaFaSpA cohort (van Hoeven et al. 2014)	>=x	x or more criteria (signs, symptoms and/or HLA-B27 positivity): • IBP • arthritis • enthesitis • dactylitis • psoriasis • IBD • uveitis • family history • good response to NSAIDs • elevated CRP • HLA-B27 positivity
Braun (2011)	>=x	x or more criteria for the recognition of axial SpA met: • age at onset ≤35 • wakening in the second half of the night • alternating buttock pain • improvement by NSAIDs within 48h • improvement by movement, not rest.
Braun (2013)	Buttock OR HLA B27	either buttock pain or HLA-B27 positivity
Braun (2013)	2-step	2 or more of the following: • improvement by movement • buttock pain • history of psoriasis or • HLA-B27 positivity

Study	Strategy	Description
Braun (2013)	>=x	x or more of the following criteria: • age at onset of chronic BP ≤35 • waking during the second half of the night • buttock pain • improvement by movement • improvement by NSAIDs within 48 h • first-grade relatives with AS • history of arthritis • history of enthesitis • history of psoriasis • HLA-B27 positivity
Sieper (2013)	as specified	specified combinations of features
HLA B27	alone	from evidence synthesis for this guideline (see @@)

4.1.2.2.2   Results

The model predicts that, on average, a person with axial spondyloarthritis who is correctly referred for specialist assessment at their first contact with healthcare services accrues just over 1 QALY more, over their lifetime, than a similar person who is not referred. However, timely referral is also estimated slightly to increase lifetime healthcare costs. This is because more people end up receiving costly interventions – notably biological DMARDs – earlier in their disease course (and remain on them for longer). This additional expense is partially offset by a reduction in background healthcare costs, with the net result that the average true-positive referral costs the NHS around 2% more, over their lifetime, than the average false negative. The costs accrued by specialist care in identifying the negative disease status of false-positive referrals is estimated at £559 each.

Table 10: Base-case deterministic cost–utility results – costs and QALYs associated with diagnostic outcomes

	Discounted lifetime costs			Discounted lifetime QALYs
	Background	**Specific**	**Total**	
True positives	£79,951	£27,356	£107,307	14.571
False negatives	£83,684	£21,282	£104,966	13.534
False positive	–	£559	£559	–
True negatives	–	–	–	–

Base-case cost–utility results are tabulated in Table 11 and illustrated in Figure 2. Results are presented for 18 of the strategies for which published data were available, as well as 2 additional scenarios – one that provides an approximation of 'current practice' and one that shows what would happen if everyone was referred to specialist care. The former is based on data on the proportion of people who are diagnosed on first presentation (NASS 2013); the latter is easily simulated with a sensitivity of 100% – that is, everyone with SpA becomes a true-positive referral – and a specificity of 0% – that is, no one who does not have SpA becomes a true-negative non-referral.

One strategy with apparently good sensitivity (>80%) and specificity (>75%) is the Braun (2013) '2-step' algorithm, in which people with possible SpA are referred on the basis of clinical questions and/or HLA-B27 positivity. However, the GDG expressed doubts about how methodologically sound, clinically meaningful and practically replicable the proposed algorithm is (especially in its reliance on reported both-sided buttock pain; see Braun 2013). For this reason, incremental results are shown for a decision-space that includes this strategy and one that excludes it.

1

Table 11: Base-case deterministic cost–utility results – possible strategies

| Strategy | Accuracy | | Costs (£) | Effects (QALYs) | Incremental | | | | | |
| | Sensitivity (%) | Specificity (%) | | | Including Braun (2013) 2-step | | | Excluding Braun (2013) 2-step | | |
					Costs (£)	Effects (QALYs)	ICER (£/QALY)	Costs (£)	Effects (QALYs)	ICER (£/QALY)
'Current practice'	10.7%	99.4%	£5,264	0.6823						
Braun (2013): >=5	53.3%	95.3%	£5,370	0.7043	£106	0.0221	£4,812	£106	0.0221	£4,812
Van Hoeven (SSB27): >=3	27.1%	88.9%	£5,371	0.6908	£1	-0.0136	dominated	£1	-0.0136	dominated
Braun (2011): >=4	47.8%	86.1%	£5,378	0.7015	£8	-0.0028	dominated	£8	-0.0028	dominated
Van Hoeven (2015): >=2	41.1%	82.4%	£5,390	0.6980	£20	-0.0063	dominated	£20	-0.0063	dominated
HLA B27: alone	68.3%	84.8%	£5,439	0.7121	£69	0.0078	ext. dom.	£69	0.0078	£8,943
Braun (2013): 2-step	80.4%	75.4%	£5,495	0.7184	£125	0.0140	£8,876	–	–	–
Van Hoeven (SSB27): >=2	64.1%	66.1%	£5,528	0.7099	£33	-0.0084	dominated	£88	-0.0022	dominated
Van Hoeven (2015): >=1.5	74.7%	57.6%	£5,561	0.7155	£66	-0.0029	dominated	£121	0.0034	ext. dom.
Braun (2013): >=4	86.0%	63.4%	£5,567	0.7213	£72	0.0029	£24,747	£127	0.0092	£13,839
Braun (2011): >=3	78.8%	46.4%	£5,625	0.7175	£58	-0.0037	dominated	£58	-0.0037	dominated
Braun (2013): Buttock OR HLA B27	89.7%	40.3%	£5,677	0.7232	£110	0.0019	ext. dom.	£110	0.0019	ext. dom.
Van Hoeven (2015): >=1.0	92.6%	39.0%	£5,680	0.7247	£114	0.0034	ext. dom.	£114	0.0034	ext. dom.
Van Hoeven (SSB27): >=1	99.7%	28.9%	£5,753	0.7284	£187	0.0071	£26,199	£187	0.0071	£26,199
Braun (2013): >=3	93.5%	26.7%	£5,757	0.7252	£4	-0.0032	dominated	£4	-0.0032	dominated
Braun (2011): >=2	96.5%	17.2%	£5,801	0.7267	£48	-0.0017	dominated	£48	-0.0017	dominated
van Hoeven (ASAS): >=1	99.7%	18.6%	£5,805	0.7284	£52	0.0000	ext. dom.	£52	0.0000	ext. dom.
Braun (2013): >=2	97.2%	7.3%	£5,857	0.7271	£104	-0.0013	dominated	£104	-0.0013	dominated
Braun (2013): >=1	99.1%	2.6%	£5,883	0.7281	£130	-0.0003	dominated	£130	-0.0003	dominated
'Refer everybody'	100.0%	0.0%	£5,896	0.7286	£143	0.0001	£992,832	£143	0.0001	£992,832

ext. dom. = extendedly dominated

2

(1) Van Hoeven (SSB27): >=1
(2) Van Hoeven (SSB27): >=2
(3) Van Hoeven (SSB27): >=3
(4) Braun (2013): >=1
(5) Braun (2013): >=2
(6) Braun (2013): >=3
(7) Braun (2013): >=4

(8) Braun (2013): >=5
(9) Van Hoeven (2015): >=1.0
(10) Van Hoeven (2015): >=1.5
(11) Van Hoeven (2015): >=2
(12) Braun (2013): Buttock OR HLA B27
(13) Braun (2013): 2-step
(14) Braun (2011): >=2

(15) Braun (2011): >=3
(16) Braun (2011): >=4
(17) van Hoeven (ASAS): >=1
(18) HLA B27: alone
(19) 'Refer everybody'
(20) 'Current practice'

Red continuous line shows cost-effectiveness frontier excluding Braun (2013) 2-step; blue dotted line shows frontier including Braun (2013) '2-step'
Dashed grey lines in background indicate ICER gradients of £20,000/QALY

Figure 2: Base-case deterministic results – cost–utility plane

QALY gains appear small, in absolute terms; however, it should be remembered that a substantial majority (95%, in the base case) of simulated patients in the model do not have SpA and, thus, experience no benefit or harm from better or worse recognition of SpA. This means that the substantial gains in quality of life for the minority of people who do have SpA appear, on face value, to be diluted by the experience of people who do not have disease. For example, under the Braun (2013) >=4 strategy, the average person with SpA gains 0.781 QALYs compared with estimated current practice. However, it is necessary to account for people without disease in the denominator of cost-per-QALY calculations, as the costs they incur are important constituents of the numerator.

If it is considered credible that the reported results of the Braun (2013) '2-step' algorithm can be replicated in NHS practice, then it is likely to be judged the optimal strategy. Compared with approximated current practice, it produces 0.036 QALYs per person at an incremental cost of £231. Several strategies have somewhat superior sensitivity and, as a consequence, somewhat superior effectiveness but, because all these strategies are also less specific than the '2-step', the incremental benefit they provide comes at an additional cost that exceeds £20,000 per QALY gained. For example, Braun (2013) >=4, being 5% more sensitive than the '2-step', is associated with 0.003 extra QALYs (approximately 1 quality-adjusted life-day) but, because it is 12% less specific, it also costs £72 per presenting person more. This produces an ICER of £24,750 per QALY gained.

1   If the Braun (2013) '2-step' algorithm is excluded from the decision space, Braun (2013) >=4
2   is likely to be considered to represent the best balance of costs and benefits. Compared with
3   approximated current practice, it produces 0.039 QALYs per presenting person at an
4   incremental cost of around £300 and, in incremental analysis, it is associated with an ICER
5   of £13,800 compared with the next-cheapest non-dominated alternative (HLA-B27 alone).
6   Again, slightly more QALYs may be gained by other strategies; in this case, van Hoeven
7   (SSB27) >=1 (which is 100% sensitive but only 29% specific) generates 0.007 extra QALYs
8   at an incremental cost approaching £200 per presenting case, leading to an ICER of £26,200
9   per QALY gained compared with Braun (2013) >=4.

10   Outputs of probabilistic sensitivity analysis (omitting Braun [2013] '2-step') are consistent with
11   the deterministic base case (see Figure 3). If QALYs are valued at £20,000 each, there is a
12   99.9% probability that one of the referral strategies simulated represents better value for
13   money than current practice. The probability that Braun (2013) >=4 is optimal at that
14   threshold is 39%.

15

Bold line indicates cost-effectiveness acceptability frontier
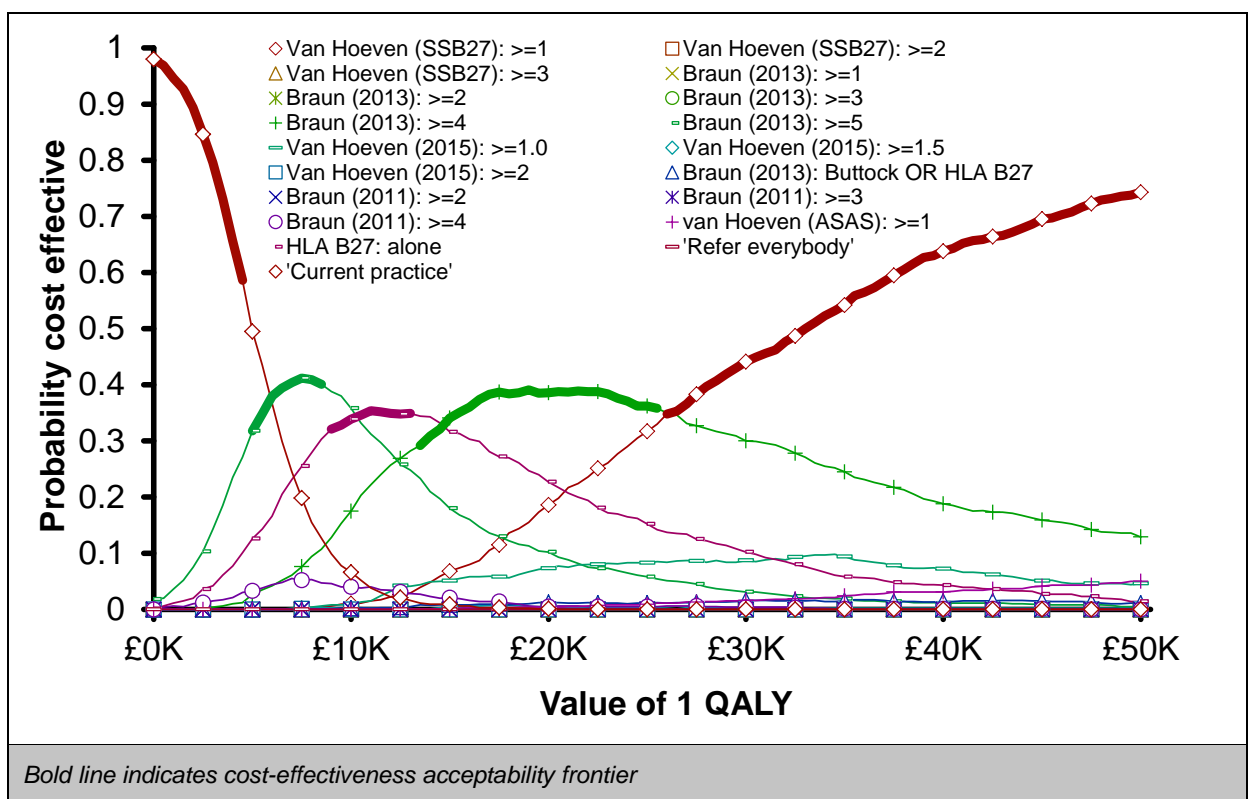
16   **Figure 3: Probabilistic sensitivity analysis – cost-effectiveness acceptability curve and**
17   **frontier**

18   One-way sensitivity analysis shows that:

19   • Braun (2013) >=4 would be preferred to Braun (2013) '2-step' with plausible alterations to
20   several parameters, including

21   o if true prevalence was above 6% (base case 5%)

22   o if sensitivity of Braun (2013) '2-step' was less than 79% (base case 80.4%)

23   o if specificity of Braun (2013) '2-step' was less than 72.5% (base case 75.4%)

24   • Braun (2013) >=4 would only represent poor value for money compared with current
25   practice if true prevalence was 1.5% or lower (base case 5%)

26   • Van Hoeven (2014) >=1 would be preferred to Braun (2013) >=4 with plausible alterations
27   to some parameters, including

1   o if true prevalence was above 6.5% (base case 5%)

2   o if average BASFI scores were assumed to be as high as 7 at the start of biological
3   DMARD therapy (base case 5.3 [AS] / 4.9 [nrAxSpA])

4.1.3   Evidence statements

4.1.3.1   Signs, symptoms and risk factors

4.1.3.1.1   People presenting with axial symptoms

Individual factors that INCREASE the probability that a person presenting with axial symptoms has spondyloarthritis

On their own, the following factors increase the probability that a person presenting with axial symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase in risk:

- Low-quality evidence:
  - Dactylitis (4 studies; total n=1,785).

On their own, the following factors increase the probability that a person presenting with axial symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** or **large** increase in risk:

- Low-quality evidence:
  - Uveitis (4 studies; total n=1,914).

On their own, the following factors increase the probability that a person presenting with axial symptoms has spondyloarthritis to a **slight** degree:

- High-quality evidence:
  - A family history of spondyloarthritis (6 studies; total n=2,908).
  - Age 35 or under at onset of back pain (in people aged 45 or under at onset of back pain) (1 study; total n=322).
  - Age 40 or under at onset of back pain (in people aged 45 or under at onset of back pain) (1 study; total n=649).
- Moderate-quality evidence:
  - Good response to NSAIDs (7 studies; total n=3,145).
- Low-quality evidence:
  - Inflammatory back pain (as defined by ASAS criteria) (4 studies; total n=1,776).
- Very low-quality evidence:
  - Inflammatory back pain (as defined by Calin criteria) (3 studies; total n=1,105).

On their own, the following factors increase the probability that a person presenting with axial symptoms has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** increase in risk:

- Low-quality evidence:
  - Buttock pain (4 studies; total n=1,951).
- Very low-quality evidence:
  - Absence of neck pain (1 study; total n=92).

Individual factors that DECREASE the probability that a person presenting with axial symptoms has spondyloarthritis

On their own, the following factors decrease the probability that a person presenting with axial symptoms has spondyloarthritis to a degree that is most likely to be **large**; however, at a 95% confidence level, data are also consistent with a **slight**, **moderate** or **very large** decrease in risk:

- Very low-quality evidence:
  - Neck pain (1 study; total n=92).

On their own, the following factors decrease the probability that a person presenting with axial symptoms has spondyloarthritis to a **moderate** degree:

- Low-quality evidence:
  - Absence of inflammatory back pain (as defined by Calin criteria) (3 studies; total n=1,105).

On their own, the following factors decrease the probability that a person presenting with axial symptoms has spondyloarthritis to a **slight** degree:

- High-quality evidence:
  - Absence of family history of spondyloarthritis (6 studies; total n=2,908).
  - Absence of inflammatory back pain (as defined by Berlin criteria) (2 studies; total n=1,013).
- Moderate-quality evidence:
  - Absence of dactylitis (4 studies; total n=1,785).
  - Heel enthesitis (2 studies; total n=1,357).
- Low-quality evidence:
  - Absence of buttock pain (4 studies; total n=1,951).
  - Absence of uveitis (5 studies; total n=2,125).

On their own, the following factors decrease the probability that a person presenting with axial symptoms has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in risk:

- Moderate-quality evidence:
  - Age 36–45 at onset of back pain (1 study; total n=322).
  - Age 41–45 at onset of back pain (1 study; total n=649).
- Low-quality evidence:
  - Absence of a good response to NSAIDs (7 studies; total n=3,145).
- Very low-quality evidence:
  - Absence of inflammatory back pain (as defined by ASAS criteria) (4 studies; total n=1,776).

4.1.3.1.2 People presenting with peripheral symptoms

Individual factors that INCREASE the probability that a person presenting with peripheral symptoms has spondyloarthritis

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **moderate** or **large** increase in risk:

- Low-quality evidence:
  - Oligoarthritis (2 studies; total n=299).

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **slight, moderate** or **large** increase in risk:

- Low-quality evidence:
  - o Inflammatory back pain (as defined by Calin criteria) (1 study; total n=81).

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **large**; however, at a 95% confidence level, data are also consistent with a **slight**, **moderate** or **very large** increase in risk:

- Very low-quality evidence:
  - o Dactylitis (2 studies; total n=229).

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a **moderate** degree:

- Moderate-quality evidence:
  - o Arthritis (1 study; total n=191).

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase in risk:

- Moderate-quality evidence:
  - o Heel enthesitis (1 study; total n=266).

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight**, **large** or **very large** increase in risk:

- Very low-quality evidence:
  - o Presence of a preceding infection (2 studies; total n=638).

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a **slight** degree:

- Low-quality evidence:
  - o A family history of psoriasis (2 studies; total n=1,909).

On their own, the following factors increase the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** increase in risk:

- Low-quality evidence:
  - o Uveitis (4 studies; total n=827).
- Very low-quality evidence:
  - o Nail disease (5 studies; total n=3,568).

Individual factors that DECREASE the probability that a person presenting with peripheral symptoms has spondyloarthritis

On their own, the following factors decrease the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **moderate** or **large** decrease in risk:

- Absence of arthritis (1 study; total n=191).

On their own, the following factors decrease the probability that a person presenting with peripheral symptoms has spondyloarthritis to a **slight** degree:
- High-quality evidence:
  - Absence of heel enthesitis (1 study; total n=266).
- Low-quality evidence:
  - Absence of family history of psoriasis (2 studies; total n=1,909).
  - Absence of family history of spondyloarthritis (2 studies; total n=666).
  - Absence of nail disease (5 studies; total n=3,568).
- Very low-quality evidence:
  - Absence of oligoarthritis (2 studies; total n=299).

On their own, the following factors decrease the probability that a person presenting with peripheral symptoms has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in risk:
- Moderate-quality evidence:
  - Absence of inflammatory back pain (as defined by Calin criteria) (1 study; total n=81).

Individual factors that DO NOT CLEARLY ALTER the probability that a person presenting with peripheral symptoms has spondyloarthritis

On their own, the following factors do not clearly alter the probability that a person presenting with peripheral symptoms has spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in risk:
- High-quality evidence:
  - Absence of inflammatory back pain (as defined by ad hoc or unreported criteria) (1 study; total n=266).
- Moderate-quality evidence:
  - Absence of nonspecific back pain (1 study; total n=372).
  - Fatigue (2 studies; total n=329).
  - Inflammatory back pain (as defined by ad hoc or unreported criteria) (1 study; total n=266).
- Low-quality evidence:
  - A family history of spondyloarthritis (2 studies; total n=666).
  - Absence of a preceding infection (2 studies; total n=638).
  - Absence of dactylitis (2 studies; total n=229).
  - Absence of fatigue (2 studies; total n=329).
  - Absence of uveitis (4 studies; total n=827).
  - Enthesitis (4 studies; total n=867).
  - Nonspecific back pain (1 study; total n=372).
- Very low-quality evidence:
  - Absence of enthesitis (4 studies; total n=867).

4.1.3.1.3 All spondyloarthritis

Individual factors that INCREASE the probability that a person has spondyloarthritis

On their own, the following factors increase the probability that a person has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase in risk:

- Low-quality evidence:
  - Uveitis (11 studies; total n=3,887).

On their own, the following factors increase the probability that a person has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** or **large** increase in risk:

- Low-quality evidence:
  - Presence of a preceding infection (7 studies; total n=2,817).
- Very low-quality evidence:
  - Dactylitis (8 studies; total n=2,888).

On their own, the following factors increase the probability that a person has spondyloarthritis to a **slight** degree:

- Moderate-quality evidence:
  - Good response to NSAIDs (9 studies; total n=4,019).
  - Inflammatory back pain (as defined by ad hoc or unreported criteria) (6 studies; total n=3,253).
- Low-quality evidence:
  - Inflammatory back pain (as defined by Calin criteria) (5 studies; total n=1,285).

On their own, the following factors increase the probability that a person has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** increase in risk:

- Moderate-quality evidence:
  - A family history of spondyloarthritis (12 studies; total n=5,395).
  - Buttock pain (6 studies; total n=2,770).
- Very low-quality evidence:
  - Psoriasis (from 9 studies; total n=4,187).

Individual factors that DECREASE the probability that a person has spondyloarthritis

On their own, the following factors decrease the probability that a person has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** decrease in risk:

- Low-quality evidence:
  - Absence of inflammatory back pain (as defined by Calin criteria) (5 studies; total n=1,285).

On their own, the following factors decrease the probability that a person has spondyloarthritis to a **slight** degree:

- High-quality evidence:
  - Absence of buttock pain (6 studies; total n=2,770).
- Moderate-quality evidence:
  - Absence of a good response to NSAIDs (9 studies; total n=4,019).

- o Absence of a preceding infection (7 studies; total n=2,817).
- o Absence of dactylitis (8 studies; total n=2,888).
- o Absence of uveitis (11 studies; total n=3,887).
- Low-quality evidence:
  - o Absence of a family history of spondyloarthritis (12 studies; total n=5,395).
  - o Absence of enthesitis (14 studies; total n=4,797).
  - o Absence of psoriasis (9 studies; total n=4,187).

On their own, the following factors decrease the probability that a person has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in risk:

- Low-quality evidence:
  - o Absence of inflammatory back pain (as defined by ad hoc or unreported criteria) (6 studies; total n=3,253).

Individual factors that DO NOT CLEARLY ALTER the probability that a person has spondyloarthritis

On their own, the following factors do not clearly alter the probability that a person has spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in risk:

- Low-quality evidence:
  - o Absence of arthritis (9 studies; total n=3,735).
  - o Absence of fatigue (2 studies; total n=329).
  - o Absence of heel enthesitis (5 studies; total n=3,185).
  - o Enthesitis (14 studies; total n=4,797).
  - o Presence or absence of inflammatory bowel disease (7 studies; total n=3,790).
  - o Presence or absence of nonspecific back pain (in people with other presenting complaints) (3 studies; total n=1,248).
- Very low-quality evidence:
  - o Arthritis (9 studies; total n=3,735).
  - o Heel enthesitis (5 studies; total n=3,185).
  - o Presence or absence of morning stiffness (2 studies; total n=1,109).

4.1.3.2 Indicators for referral

Low- to moderate-quality evidence from 4 studies reporting the evaluation of referral algorithms, based on 2 underlying cohort studies, found different referral strategies gave a wide range of different sensitivities and specificities for the identification of axial spondyloarthritis. These referral strategies were evaluated in the economic modelling undertaken as part of this guideline (see section 2.1.2.2 and Appendix H).

4.1.3.3 Comparison of referral strategies

Moderate-quality evidence from 2 studies detected no differences in the proportion of correct diagnoses in those referred following simple vs complex referral strategies.

4.1.3.4 Case-finding for spondyloarthritis in people with acute anterior uveitis

CONFIRMING that a person with acute anterior uveitis has spondyloarthritis

If a person with acute anterior uveitis **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a **very large** degree:

- Moderate-quality evidence:
    - o DUET algorithm (2 studies; total n=173).

EXCLUDING the possibility that a person with acute anterior uveitis has spondyloarthritis

If a person with acute anterior uveitis **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **large** decrease in probability:

- Moderate-quality evidence:
    - o DUET algorithm (2 studies; total n=173).

4.1.3.5 Delays to diagnosis

4.1.3.5.1 Quantitative evidence

Gender

Very low-quality evidence from 5 studies found no significant difference in diagnostic delay between males and females.

HLA B27

Very-low quality evidence from 4 studies evaluated diagnostic delays in people who were HLA B27 negative compared with those who were HLA B27 positive, with 2 studies finding significantly increased diagnostic delays and 2 studies finding no significant difference.

Peripheral joint involvement

Very low-quality evidence from 3 studies found no significant difference in diagnostic delay between people with and without peripheral joint involvement.

Inflammatory back pain

Very low-quality evidence from 3 studies evaluated diagnostic delays in people without inflammatory back pain compared with those with inflammatory back pain, with 2 studies finding significantly increased diagnostic delays and 1 study finding no significant difference.

Extra-articular involvement

Very low-quality evidence from 1 study evaluated diagnostic delays in people with extra-articular involvement compared with those without extra-articular involvement, with 1 study finding significantly increased diagnostic delays and 1 study finding no significant difference

Family history

Very low-quality evidence from 1 study evaluated diagnostic delays in people without a first-degree relative with spondyloarthritis compared with those with, with 1 study finding significantly increased diagnostic delays and 3 studies finding no significant difference

Age

Very low-quality evidence in 2 studies found significantly increased diagnostic delay in people younger than 16 compared with those who were older.

Very low-quality evidence in 1 study found no significant difference in diagnostic delay between those who were older or younger than 17.

Very low-quality evidence from 1 study found that increasing age at diagnosis in adults was significantly associated with an increased mean diagnostic delay.

GP knowledge, belief and experiences

Very low-quality evidence from 1 study found that between 13% and 90% of GPs were able to identify individual symptoms associated with inflammatory back pain.

4.1.3.5.2 Qualitative evidence

Journey to diagnosis

Very low-quality evidence from 1 study identified the following themes in the journey to diagnosis of those with spondyloarthritis: being unsure when symptoms started of what was going on, feeling like they had to fight for a diagnosis, being adrift where diagnosis was delayed, and that diagnosis provided a sense of relief and represented the start of a journey.

GP knowledge, belief and experiences

Very low-quality evidence from 1 study identified the following themes from interviews with a small number of GPs: knowledge gaps in differentiating mechanical back pain from inflammatory back pain and in the ability to describe axial spondyloarthritis, that ankylosing spondylitis was considered to be almost exclusively diagnosed in men, that delay in diagnosis was a concern linked to both patients' and doctors' delay, some awareness of extra-articular manifestations.

4.1.3.6 Health economic evidence

A directly applicable original cost–utility analysis with minor limitations developed for this guideline explored potential referral rules for people presenting with back pain. Results suggest that many potential strategies would improve quality-adjusted life expectation at reasonable cost, compared with current practice. Assuming QALYs are valued at £20,000 each, the optimal approach is likely to be either a 2-step strategy (in which people are referred if they have **either** 2 or more of improvement by movement, buttock pain or a history of psoriasis **or** they are HLA-B27 positive) or, if that strategy is removed from the decision space, referring people who have 4 or more features from a list of 10 (age at onset ≤35; waking during the second half of the night; buttock pain; improvement by movement; improvement by NSAIDs within 48 hours; first-grade relatives with ankylosing spondylitis; a history of arthritis; a history of enthesitis; a history of psoriasis; HLA-B27 positivity). Model outputs are critically dependent on the assumed true prevalence of axial spondyloarthritis among people presenting with chronic back pain that started at age 45 or younger.

4.1.4 Evidence to recommendations

Relative value of different indications and outcomes	Signs, symptoms, risk factors and referral strategies
	The GDG agreed that, for all questions with diagnostic accuracy data, the most useful measures were positive and negative likelihood ratios. A positive likelihood ratio greater than 2 and a negative likelihood ratio less than 0.5 were considered to be clinically important results, in line with the suggestions of Jaeschke et al.

(1994).

The GDG noted the importance of having a recommendation to emphasise the diversity in presentation of spondyloarthritis, making specific reference to the signs and symptoms that demonstrate a statistically significant positive likelihood ratio. The GDG wants the guideline to raise awareness of spondyloarthritis amongst non-specialists and therefore see value in highlighting signs and symptoms to look out for in addition to the more specific recommendations around the particular combinations of features that should lead to specialist referral. This may serve to assist in cases where the person presenting meets insufficient criteria to be referred on that occasion, but may benefit from being followed up over time for the emergence of further signs and symptoms.

When recommending referral criteria, the key factors relate to the ability of any strategy not to miss people with spondyloarthritis (highly sensitive) whilst not overburdening rheumatology services with too many people who do not have the condition (highly specific). A strategy which is inclusive enough not to miss anyone with the condition will be less specific; establishing referral criteria which ensure that every person with spondyloarthritis is captured inevitably means more people without the condition are referred inappropriately. Whilst taking this into consideration, the GDG expressed a preference for potentially being over-inclusive as a means of ameliorating the current problem of delayed referral and under-diagnosis. This problem is particularly marked in axial spondyloarthritis.

The GDG accepted that individual signs, symptoms and risk factors might not provide enough information on their own to warrant a referral or make a correct diagnosis in either axial or peripheral spondyloarthritis

Obstacles to diagnosis

Spondyloarthritis is considered to be an under-diagnosed condition; those who have been diagnosed with spondyloarthritis may experience a substantial delay in the time taken to reach that diagnosis.

The GDG agreed that it was important to consider if there are obstacles to achieving a prompt diagnosis that could be identified.

The GDG agreed a list of potential barriers that were included in the review protocol but did not consider this to be exhaustive list and recognised that other obstacles may be identified.

Trade-off between benefits and harms	The delay to diagnosis of spondyloarthritis, particularly in axial disease, makes a missed diagnosis of utmost importance to the GDG. Therefore sensitivity was seen to be of high importance when recommending a strategy for referral. However, the GDG remained aware of the importance of specificity in any referral strategy as there exists a point at which the impact of over-referral on specialist services would become too great, possibly resulting in further delay in diagnosis due to lack of resources to meet demand.
	The GDG agreed that special emphasis should be placed on immediate referral to an ophthalmologist where acute anterior uveitis is suspected. Failure to recognise and appropriately treat this symptom can result in sight-loss.
	Urgent referral is also recommended for people presenting with suspected new-onset inflammatory synovitis, because delays in appropriate intervention can lead to permanent joint damage, with long-term health consequences, and these people are highly likely to be on an early inflammatory arthritis pathway (either spondyloarthritis or rheumatoid arthritis). The GDG consider that only inflammatory polyarthritis was likely to be indicative of spondyloarthritis, and restricted their focus to that, as people with fewer joints affected

would be more likely to be referred to specialist care for assessment for rheumatoid arthritis.

Obstacles to diagnosis

The GDG discussed how traditional beliefs about how people with spondyloarthritis will present to primary care affected appropriate diagnosis. They considered this to be particularly pertinent for axial spondyloarthritis where there may be an over-reliance on assuming that those with axial spondyloarthritis present with lower back pain. Consequently the GDG agreed a recommendation that highlights the importance of recognising misconceptions about how axial spondyloarthritis presents and that presentation may be heterogeneous. The GDG agreed and acknowledged that there may be negative consequences related to delayed diagnosis for people with axial spondyloarthritis.

The GDG agreed that the evidence presented did not find that gender had an impact on delay to diagnosis, although it did suggest that GPs perceive ankylosing spondylitis to be a predominantly male condition. The GDG noted the small number of women in the included studies and the overall very low quality of the evidence. The traditional belief in practice that axial spondyloarthritis is more prevalent in men was further discussed and the GDG agreed that a recommendation raising awareness that this belief is incorrect is warranted. As evidence for this, the GDG noted that large cross-sectional studies have found that equal numbers of women and men are diagnosed as having axial spondyloarthritis (ASAS 53% male; DESIR 50% male; SPACE 45% male).

The GDG noted that the included evidence did not include participants who had been referred from non-primary care settings to rheumatology, such as from dermatology or physiotherapy. The GDG discussed the variety of symptom presentations in peripheral spondyloarthritis and considered that a recommendation to raise awareness of this was important and that this should include the possibility that spondyloarthritis can present prior to the associated condition.

Economic considerations	The health economic model for axial disease aimed to represent the difference in the quality of life and the costs to the NHS between people with spondyloarthritis who receive a positive diagnosis and those who remain undiagnosed. This, along with the costs of correcting a false negative diagnosis, enables the quantification of the relative value of sensitivity and specificity and thus the trade-offs between benefits and harms associated with the accuracy of potential referral strategies.

Economic modelling conducted for this review question was confined to the axial manifestation of the condition as this is where the greatest burden of under-diagnosis is thought to occur.

The economic model demonstrates that sensitivity largely drives the cost effectiveness of the strategies but that, once this metric reaches reasonably high levels, specificity also plays a role.

The prevalence estimate of axial spondyloarthritis among people presenting to primary care with low back pain is key to the cost effectiveness of the strategies modelled. An evidence-based estimate of this which the GDG felt could be applied to the population in question does not exist; therefore the GDG suggested using a 5% value in the base case. Prevalence does not affect the relative ranking of each of the potential referral strategies in terms of costs and effects individually: strategies that generate more QALYs will always generate more QALYs and strategies that cost more will always cost more regardless of prevalence. As a result, when the prevalence parameter is altered in the model, the same strategies lie |

on the frontier on the cost–utility plane (meaning that they offer a health benefit but at an increased cost). However, the balance between costs and benefits, and whether the strategies offer a reasonable use of NHS resources with ICERs below £20,000 per QALY gained is heavily influenced by the prevalence.

The optimum referral strategy in the health economic analysis is one which offers the most additional health at an acceptable cost per QALY. In the model's base case, the strategy that appears optimal is a 2-step approach which involves referral in the presence of 2 out of the 3 following signs and symptoms:

- Both-sided buttock pain

- Psoriasis

- Improvement of back pain by movement

If fewer than 2 of these factors are met then an HLA B27 test is to be conducted and a referral made if the test result is positive.

When the GDG discussed this strategy, it raised both clinical and methodological concerns:

- The GDG stated that the prominence of both-sided buttock pain in the algorithm did not accord with their experience, or with other evidence about this symptom. The GDG acknowledged that the presence of both-sided buttock pain is a symptom which, as an individual factor, has some usefulness in ruling people in, but the lack of the symptom does not confirm the absence of disease. Moreover, the GDG agreed that, even if the diagnostic accuracy of the strategy were to be proven in a validation cohort, its usefulness in general practice is limited due to ambiguity in the way in which buttock-pain is defined and the difficulty that some patients may have in distinguishing buttock pain from low back pain. The GC expressed the view that, while considering buttock pain when there is a suspicion of inflammatory back pain is useful in a specialist environment, out of this setting, the investigation may yield less accurate results. Therefore, use of the 2-step recommendation in non-specialist settings may not aid differential diagnosis, leading to inappropriate identification of sciatica or fibromyalgia as buttock pain indicating suspected spondyloarthritis.

- The strategy is the best performing strategy within the cohort in which it was derived; however, it has not been externally validated. The authors considered a number of other prescriptive strategies (not reported) and highlighted the one they considered to have the optimum balance of sensitivity and specificity. This means the performance of the 2-step strategy, in terms of its ability to appropriately distinguish between people with and without spondyloarthritis, may reflect the chance characteristics of the studied cohort, and may not generalise to other populations presenting with signs and symptoms suggestive of spondyloarthritis.

For these reasons, the GDG concluded that there was too much risk that the results reported for the 2-step strategy would not be replicated in practice. Therefore, the approach could not be safely recommended, and the GDG requested that it should be excluded from the decision-space and the analysis recalculated.

When the analysis was revised in this way, the optimal strategy was Braun>=4.The diagnostic accuracy is evaluated of strategies with varying cut-offs according to the number of factors (out of 10: age at onset of chronic BP of under 35 years, waking during the second half of the night due to back pain, buttock pain, improvement by movement and not by rest, improvement by NSAIDs within 48h, first-grade relatives with AS, history of arthritis, history of enthesitis, history of psoriasis). All possible thresholds were considered in the

original model, and the optimal cut-off was a rule that refers people with 4 or more of the relevant features.

It was noted that evidence for this strategy comes from the same study as the 2-step strategy. However, as a broad range of factors are considered, and no individual feature is specified as critical to the referral decision, the findings are not deemed to suffer from the same degree of susceptibility to chance findings as the 2-step strategy. This also has the issue of not having been externally validated as yet, but the GDG were confident that the range of features included in the strategy meant there would be less likelihood of inappropriate non-referral than with the 2-step strategy.

It was noted that, in contrast to most other potential referral schemas, which include inflammatory back pain as a single item, the Braun>=4 strategy separates it into its constituent elements, and considers each of these as a discrete marker that raises the chance of spondyloarthritis.

The GDG agreed that this is an advantageous approach, for at least 2 reasons: (a) there are many different definitions of inflammatory back pain, and it is unclear that these have equivalent diagnostic accuracy (see 4.1.3.1); therefore, it is helpful to bypass this complexity by specifying the individual features of note; (b) it makes the strategy more user-friendly for people who are unfamiliar with the signs and symptoms associated with the presentation of inflammatory back pain.

The GDG discussed the potential of this strategy to exclude groups of individuals presenting with an insufficient number of parameters from being referred. Theoretically, a person with a positive HLA B27 status and two other factors (from inflammatory back pain, arthritis, enthesitis (heel), uveitis, dactylitis, psoriasis, Crohn's/colitis, good response to NSAIDs, family history for SpA, elevated CRP) would warrant a final diagnosis of non-radiographic axial SpA according to the ASAS classification criteria. The GDG noted that it was unlikely that people with spondyloarthritis and a positive HLA B27 status would present with only 2 of the features from the Braun >=4 strategy list, especially since it decomposes inflammatory back pain into its constituent elements; therefore, the GDG was not unduly concerned about this possibility.

An argument could be made that the next strategy on the cost–utility frontier (van Hoeven SSB27 >=1) could also be a viable approach to adopt. This would be especially true if prevalence of axial spondyloarthritis in people presenting with low-back pain were believed to be only slightly higher than the base-case value of 5%. If prevalence were greater than 6.5%, van Hoeven [SSB27 >=1] would become the optimal strategy, with an ICER less than £20,000 per QALY compared with Braun >=4. However, the GDG thought this strategy was unlikely to provide a viable approach. Although it is 100% sensitive (meaning no one with axial spondyloarthritis would be missed), this benefit comes at the expense of much-reduced specificity, which means that many more false-positive cases would result. For example, if true prevalence were 7%, fewer than 1 in 10 people referred under this rule would ultimately receive a diagnosis of spondyloarthritis. The economic analysis takes into account the cost per patient of a referral to secondary care and the average diagnostic work-up necessary in order to rule the condition in or out, so this might still theoretically be an optimal approach (as reflected in the ICER). Practically speaking, however, there is undoubtedly a tipping-point at which the current provision of rheumatology services within the NHS would have to be significantly altered in order to meet demand in terms of the number of people referred. For this reason – and because of their acknowledged uncertainty about the true prevalence of disease – the GDG agreed that it was preferable to

	recommend a strategy that (a) appeared optimal when adopting best possible parameters, and (b) reflected a balance of sensitivity and specificity that would identify the substantial majority of cases without running the risk of overburdening the service.
	Despite this, the GDG noted that false-positive referrals are, in reality, not entirely without value (as the model effectively assumes). There may be additional benefits of a consultation with a rheumatologist such as a diagnosis of peripheral disease, differential diagnosis of an alternative rheumatological condition or excluding a diagnosis of spondyloarthritis. These additional benefits have not been represented within the health economic model but could be important for both the patient and the health service as a whole.
Quality of evidence	**Signs, symptoms and risk factors**
	The following signs, symptoms and risk factors had evidence that gave rise to statistically significant positive likelihood ratios in axial and/or peripheral populations and/or across all presentations: inflammatory back pain (variously defined), enthesitis, dactylitis, inflammatory arthritis (excluding gout/rheumatoid arthritis), chronic lower back pain with earlier age of onset (under 45, under 40 and under 35, in different studies), pain that improves with NSAID use, uveitis, psoriasis and/or psoriatic nail symptoms, current or recent genitourinary infection, family history of spondyloarthritis or psoriasis.
	The GDG judged that this evidence was of sufficient relevance that these items should appear on a broad list of spondyloarthritis-related features of both axial and peripheral spondyloarthritis that should be recognised as such in specialist and non-specialist settings. Overall the GDG considered that each alone was insufficient to act as a sole referral or diagnostic criterion, with the exception of dactylitis which would need referral to a rheumatologist irrespective of whether or not it was associated with spondyloarthritis. This was an acknowledgement that some of the above features had statistically, but not clinically significant positive and/or negative likelihood ratios.
	Dactylitis demonstrated a statistically significant positive likelihood ratio of 4.26 when all studies (axial, peripheral and mixed populations) were pooled, which the GDG agreed was sufficiently large to justify a referral recommendation based on the presence of this symptom alone.
	The positive and negative likelihood ratios for enthesitis as a sign of spondyloarthritis were weak. The GDG agreed that people with suspected peripheral spondyloarthritis should not be referred on the basis of enthesitis alone. However, based on their experience and expertise, the GDG drafted a recommendation for referral where people present with enthesitis in conjunction with additional qualifiers thereby creating a good balance between sensitivity and specificity.
	Referral criteria
	The systematic review for referral criteria found 2 moderate-quality studies with head-to-head comparisons of referral strategies (Poddubnyy et al., 2011 and Sieper at al., 2013) and 2 studies (Braun et al., 2013 and van Hoeven et al., 2015) that fully evaluated the referral of people with suspected axial disease from primary to secondary care.
	Poddubnyy et al. (2011) & Sieper et al. (2013) provide additional pieces of evidence around outcomes for people who **were** referred according to specific referral strategies for axial disease, but this could not be incorporated into the economic analysis, because the same studies do not provide any follow-up of the people who were **not** referred, and this makes the reported diagnostic accuracies uninterpretable.
	The evidence from Braun et al. (2013) & van Hoeven et al. (2015) was supplemented with strategies simulated from the syntheses on

signs, symptoms, risk factors and tests based on estimates of their correlations when used in combination taken from Sieper et al. (2013), and an abstract and a letter presenting validation of a set of referral criteria which approximate the 'ASAS referral criteria' in the van Hoeven et al. (2015) cohort. The fact that these latter data were drawn from sources that are commonly considered of limited validity (a conference abstract and a letter) was not seen as a reason to dismiss or downgrade the evidence, because the methods by which the cohort in which the analyses were undertaken was recruited and analysed is well described in multiple full-length peer-reviewed publications (van Hoeven et al. 2014; van Hoeven et al. 2015). The simulated strategies allowed the best available evidence of the performance of individual factors to be combined and tested in terms of their ability to generate health at a cost acceptable to the NHS.

Moderate-quality evidence was identified for a case-finding algorithm that reliably identifies cases of spondyloarthritis in people with acute anterior uveitis. This facilitated a strong recommendation to follow this strategy.

Obstacles to diagnosis

The GDG agreed with the assessment of the evidence using the BMJ checklist and that the evidence presented was of very low quality. The GDG agreed that the included evidence was relevant to the review question, but due to concerns about the representativeness of the included samples, the small, single-centred nature of the studies and the limited detail of analysis methods reported the evidence statements could not be viewed as robust.

Initial recognition

In the initial recognition recommendation, which comes before any recommendations about referral criteria, the GDG agreed it was appropriate to adopt a lower standard of evidence for factors that would merely raise awareness that axial or peripheral spondyloarthritis was a possibility than for those which would result in referral, as no major harm would result provided it did not lead to inappropriate referrals. Therefore, the standard MCID criteria were not applied here, and all factors with statistically significant diagnostic value were considered for inclusion, even if the likelihood ratio did not meet the defined MCID threshold.

Other considerations	In identifying signs and symptoms that should prompt healthcare professionals to think of axial spondyloarthritis, the GDG agreed that, although none of the inflammatory back pain criteria stood out as being a clearly better discriminator for spondyloarthritis than others, it would prevent confusion if the recommendation contained direction to a single set of criteria to use to identify inflammatory back-pain. It chose the ASAS inflammatory back-pain criteria as these are well known amongst specialists and correspond well with clinical experience.
The evidence presented supported the GDG's experience that, outside of specialist rheumatology services, there has been a long-established notion that women and people with a negative HLA B27 status do not develop axial spondyloarthritis. The GDG therefore drafted a recommendation highlighting the equivalence of spondyloarthritis in women and men..
The GDG emphasised that spondyloarthritis can manifest in diverse ways and features may become more pronounced or identifiable with time. Therefore any referral criteria recommendations should not result in people who currently meet fewer criteria than warrant a referral being permanently ruled out from having a potential diagnosis of spondyloarthritis. There is a need for healthcare practitioners to ensure people who present with some signs, symptoms or risk factors that are suggestive of spondyloarthritis, but who do not fully

meet referral criteria at that encounter, are made aware that further symptoms can develop and to re-present for a further assessment if that becomes the case.

The GDG noted that, in addition to being a relevant risk factor for spondyloarthritis, people presenting in non-specialist settings with current acute anterior uveitis need urgent ophthalmological assessment to prevent damage to sight caused by the condition. The GDG therefore deemed it appropriate to make an urgent referral for assessment, based on the advice of the co-opted ophthalmologist as well as the wider experience of the committee.

A referral criterion for suspected axial spondyloarthritis which includes HLA B27 testing provides a mandate for this test to be used, when appropriate, in primary care, which is a departure from current practice in much of the UK. The GDG discussed the implications of this change and agreed that they were unaware of any reason to believe the test would perform differently in primary care, as the result returned by the testing laboratories is binary and should therefore be easy to interpret. Given that the economic analysis takes into consideration the costs of conducting the test as well as the accuracy of the strategy, in combination with other signs and symptoms, the GDG supported this change in practice in the interest of recommending a cost-effective referral rule. It noted that the number of tests required would be relatively small as, using the recommended referral strategy, HLA B27 testing will only be necessary to make a final decision about referral in people who present with exactly 3 of the signs or symptoms from the agreed list – people with 4 or more clinical features should be referred without further testing, and people with 2 or fewer would not meet criteria for automatic referral whether they are HLA-B27 positive or negative.

As discussed in 'Trade off of Benefits and Harms', the GDG agreed that new-onset inflammatory polyarthritis posed sufficient risk of harm if left untreated that a referral to rheumatology was warranted on the basis of this alone in people with suspected peripheral spondyloarthritis. They did not extend this recommendation to cover mono- or oligoarthritis as it was felt that people presenting with this level of joint involvement were already adequately supported by other care pathways or guidelines.

The GDG acknowledged that none of the referral strategies under consideration in the modelling offered perfect sensitivity and specificity, which means that there will be a proportion of people who will not be correctly referred/not referred. It was noted that different referral strategies relied on different criteria and it was possible that the choice of preferred criteria would have an influence on which people were incorrectly assigned to referral/non-referral. The GDG were therefore concerned that people who nearly met the selected referral criteria, but not sufficiently to be referred on that occasion, should not have spondyloarthritis prematurely ruled out. This was of particular concern given that some of the signs and symptoms listed in the referral criteria may not have occurred at initial presentation, only occurring later. The GDG therefore drafted a recommendation advising what to do when a person with suspected spondyloarthritis does not yet fully meet the referral criteria for specialist assessment, but there is still clinical suspicion of spondyloarthritis.

4.1.5 Recommendations

1.1. Do not rule out the possibility that a person has spondyloarthritis solely on the presence or absence of any individual sign, symptom or test result.

2. Suspecting spondyloarthritis

2.1. **Recognise that spondyloarthritis can have diverse symptoms and be difficult to identify, which can lead to delayed or missed diagnoses. Signs and symptoms may be musculoskeletal (for example, inflammatory back pain, enthesitis, dactylitis) or extra-articular (for example, uveitis, psoriasis [including psoriatic nail symptoms]), and risk factors include recent genitourinary infection and a family history of spondyloarthritis or psoriasis.**

2.2. **Be aware that axial spondyloarthritis:**

- affects a similar number of women as men
- can occur in people who are human leukocyte antigen B27 (HLA B27) negative
- may be present despite no evidence of sacroiliitis on a plain film X-ray.

2.3. **Be aware that peripheral spondyloarthritis may be missed, even if the onset is associated with established comorbidities (for example, psoriasis, uveitis, inflammatory bowel disease [Crohn's disease or ulcerative colitis] or a gastrointestinal or genitourinary infection).**

3. **Referral for suspected axial spondyloarthritis**

3.1. **Refer people with low back pain, that started at under 45 years and has lasted for longer than 3 months, to a rheumatologist for a spondyloarthritis assessment when at least 4 of the following are present:**

- onset of back pain at under 35 years (this further increases the likelihood that back pain is due to spondyloarthritis compared with onset of back pain at between 35 and 44 years)
- waking during the second half of the night because of symptoms
- buttock pain
- improvement with movement
- improvement within 48 hours of taking NSAIDs
- a first-degree relative with spondyloarthritis
- current or past arthritis
- current or past enthesitis
- current or past psoriasis.

If 3 criteria are present, perform an HLA-B27 test and refer to a rheumatologist for a spondyloarthritis assessment if this test is positive.

3.2. **If the person does not meet the criteria in recommendation 3.1 but clinical suspicion of axial spondyloarthritis remains, advise the person to seek repeat assessment if new signs, symptoms or risk factors listed in recommendation 3.1 develop. This may be especially appropriate if the person has current or past inflammatory bowel disease (Crohn's disease or ulcerative colitis), psoriasis or uveitis (see recommendation 5.1 for guidance on referral for immediate [same-day] ophthalmological assessment for people with acute anterior uveitis).**

4. **Referral for suspected peripheral spondyloarthritis**

4.1. **For guidance on identifying spondyloarthritis in people with an existing diagnosis of psoriasis, see the NICE guideline on psoriasis.**

4.2. **Urgently refer people with suspected new-onset inflammatory polyarthritis to a rheumatologist for a spondyloarthritis assessment, unless rheumatoid arthritis, gout or acute calcium pyrophosphate (CPP) arthritis ('pseudogout')**

is suspected. If rheumatoid arthritis is suspected, see referral for specialist treatment in the NICE guideline on rheumatoid arthritis in adults.

4.3. **Refer people with dactylitis to a rheumatologist for a spondyloarthritis assessment.**

4.4. **Refer people with enthesitis without apparent mechanical cause to a rheumatologist for a spondyloarthritis assessment if:**

- it is persistent or
- it is in multiple sites or
- any of the following are also present:
  - ο back pain without apparent mechanical cause
  - ο current or past uveitis (see recommendation 5.1 for guidance on immediate [same-day] ophthalmological assessment for people with acute anterior uveitis)
  - ο current or past psoriasis
  - ο gastrointestinal or genitourinary infection
  - ο inflammatory bowel disease (Crohn's disease or ulcerative colitis).
  - ο a first-degree relative with spondyloarthritis or psoriasis.

5. **Referral for suspected acute anterior uveitis**

5.1. **Refer people for an immediate (same-day) ophthalmological assessment if they have symptoms of acute anterior uveitis (for example, eye pain, eye redness, sensitivity to light or blurred vision).**

6. **Case-finding in people with acute anterior uveitis**

6.1. **Ophthalmologists should ask people with acute anterior uveitis whether they have:**

- consulted their GP about joint pains **or**
- experienced chronic low back pain that started at under 45 years and has lasted for longer than 3 months.

6.2. **If the person meets either of the criteria in recommendation 6.1, establish whether they have psoriasis or skin complaints that appear psoriatic on physical examination.**

- If they do, refer the person to a rheumatologist for a spondyloarthritis assessment.
- If they do not, perform an HLA-B27 test and refer the person to a rheumatologist for a spondyloarthritis assessment if this test is positive.

4.1.6 Research Recommendations

1. **What are the optimal referral criteria for people with suspected spondyloarthritis?**

Repeat the CaFaSpA study (van Hoeven et al. 2014, 2015) in a UK population. This would involve examination of GP databases to identify a cohort of people who have a diagnosis of non-specific back pain who first consulted their GP for back symptoms under the age of 45. These people would be invited for full rheumatological work-up (including: identification of signs and symptoms relevant to axial spondyloarthritis, X-ray, MRI, HLA-B27 test). All participants would be given a reference-standard diagnosis of axial spondyloarthritis or not (ideally using expert clinician opinion; if not feasible, use ASAS classification criteria). The cohort would be split into a development and validation set, to derive and validate optimal

rules for case-finding from the available data, with each candidate strategy judged according to expected cost per QALY gained (NICE model could be used to estimate these).

Why this is important

As a result of the large number of permutations of possible referral strategies, it is impractical to run separate validation studies for all referral criteria that are developed. Therefore, a single large, representative cohort study would, provided it measured the predictor variables for all reasonable referral strategies, provide the ability to develop and validate any number of possible referral strategies. The study would need to be large enough that sufficient data are available to derive new referral rules and to validate those rules in a separate, independent subset of the data. A UK-specific dataset would provide more relevant data to do this than is currently available from the Dutch CaFaSpA study. For example, that study found an HLA-B27 prevalence of 20% in people with axial spondyloarthritis and 2% in people without; much lower than estimates found elsewhere (75% and 20%, respectively). This lowers the validity of extrapolating any results found to the UK, and reinforces the need for UK-specific data to address this question.

2. **At what stage and using what criteria should people with inflammatory bowel disease be referred to a rheumatologist for a spondyloarthritis assessment?**

Why this is important

The guideline committee noted that people with inflammatory bowel disease (Crohn's disease or ulcerative colitis) are more likely to have or develop spondyloarthritis than those without. During the development of this guideline specific, validated referral rules were identified for people with inflammatory back pain or acute anterior uveitis, but not for people with inflammatory bowel disease. An inflammatory bowel disease-specific referral rule would provide additional value as the diagnostic importance of other spondyloarthritis associated features may be different in the presence of inflammatory bowel disease, something which is not possible to judge from the currently available data. There is therefore a need for the development of inflammatory bowel disease-specific referral rules, which would need to be prospectively validated in a cohort of people with confirmed inflammatory bowel disease and suspected spondyloarthritis. This study would need to follow up both those people who were and were not referred until a definitive diagnosis has been made (ideally using expert clinician opinion; if this is not feasible, using the ASAS classification criteria).

3. **What is the effectiveness and cost-effectiveness of educational interventions for healthcare professionals in order to increase the number of prompt diagnoses of spondyloarthritis?**

Why this is important

One of the major reasons identified during this guideline for the delays in diagnosis of spondyloarthritis is a lack of awareness of the condition on behalf of by healthcare professionals. This can take many forms, such as a lack of awareness of different spondyloarthritis subtypes, lack of knowledge about associated clinical features (for example, the differences between inflammatory and mechanical back pain) or characteristics of the patient populations (for example, that spondyloarthritis affects similar numbers of men and women, or that a substantial proportion of people with spondyloarthritis are HLA-B27 negative). Educational interventions to improve the level of awareness may therefore lead to reductions in diagnosis delays, but there is a lack of evidence as to the efficacy of these interventions. Randomised controlled trials of structured educational interventions are therefore needed to assess both whether they reduce the length of time it takes for people to be correctly identified and diagnosed, and whether they represent a cost-effective use of NHS resources.

4.2 Blood tests for spondyloarthritis

Review questions 7, 8 and 9

- What is the diagnostic utility of a HLA B27 test for investigating suspected spondyloarthritis?
- What is the diagnostic utility of an erythrocyte sedimentation rate test for investigating suspected spondyloarthritis?
- What is the diagnostic utility of a C-reactive protein test for investigating suspected spondyloarthritis?

4.2.1 Evidence review

The aim of this review was to assess the utility of assessing specific genetic or circulating biomarkers as part of the recognition or diagnostic process for people with suspected spondyloarthritis.

The review focussed on identifying studies that fulfilled the conditions specified in Table 12, Table 13 and Table 14.

Table 12: PICO table for question 7: HLA-B27

Population	Patients with suspected spondyloarthritis
Interventions	HLA-B27 genetic test
Comparators	Clinical opinion of spondyloarthritis was considered the preferred reference standard, with diagnosis using any specified criteria as the next preference.
Outcomes	Sensitivity Specificity Positive likelihood ratio Negative likelihood ratio Positive predictive value Negative predictive value Diagnostic odds ratio

Table 13: PICO table for question 8: ESR

Population	Patients with suspected spondyloarthritis
Interventions	Erythrocyte sedimentation rate (ESR)
Comparators	Clinical opinion of spondyloarthritis was considered the preferred reference standard, with diagnosis using any specified criteria as the next preference.
Outcomes	Sensitivity Specificity Positive likelihood ratio Negative likelihood ratio Positive predictive value Negative predictive value Diagnostic odds ratio

Table 14: PICO table for question 9: CRP

Population	Patients with suspected spondyloarthritis
Interventions	C-reactive protein (CRP)
Comparators	Clinical opinion of spondyloarthritis was considered the preferred reference standard, with diagnosis using any specified criteria as the next preference.
Outcomes	Sensitivity Specificity

	Positive likelihood ratio
	Negative likelihood ratio
	Positive predictive value
	Negative predictive value
	Diagnostic odds ratio

1   For full details of the review protocol please see Appendix C.

2   Cross-sectional studies and cohort studies were considered to be the highest quality
3   evidence available to answer this question and are graded as high in a GRADE framework if
4   conducted and reported well.

5   Systematic searches for this question identified 9,176 references, which were screened on
6   their titles and abstracts. Between papers identified from this search and papers identified
7   from other diagnostic searches in this guideline which contained relevant data, a total of 80
8   studies were retrieved for full text review. Forty of these studies were excluded as they did
9   not meet the eligibility criteria such as inappropriate study design (e.g. case reports, case
10 series with fewer than ten cases) or non-primary studies (e.g. systematic review, editorial).
11 Detailed lists of excluded studies and reasons for their exclusion are provided in Appendix F.

12 **4.2.1.1   Description of included studies**

13 Thirty cross sectional studies were identified which contained relevant information on the
14 diagnostic value of HLA-B27 testing (13 in axial spondyloarthritis, 7 in peripheral
15 spondyloarthritis and 10 in a mixed population containing people with both axial and
16 peripheral spondyloarthritis).

17 Two cross sectional studies were identified which contained relevant information on the
18 diagnostic value of ESR testing (1 in axial spondyloarthritis and 1 in a mixed population
19 containing people with both axial and peripheral spondyloarthritis).

20 Eight cross sectional studies were identified which contained relevant information on the
21 diagnostic value of CRP testing (5 in axial spondyloarthritis, 2 in peripheral spondyloarthritis
22 and 11 in a mixed population containing people with both axial and peripheral
23 spondyloarthritis).

24 Evidence tables for included studies can be found in Appendix E, with GRADE profiles
25 reported in Appendix G.

26 **4.2.1.2   Variations from protocol**

27 A specific search was conducted for these questions, to identify studies which sought to
28 evaluate the diagnostic utility of these biomarkers. However, in the course of conducting
29 other diagnostic utility questions for this guideline, further data was identified which was
30 contained incidentally in other studies. Across all of the searches conducted for any of the
31 diagnostic questions, any relevant data was extracted, if the study met the eligibility criteria,
32 regardless of whether it was identified in the specific search for this study.

33 **4.2.1.3   Minimal clinically important differences**

34 Minimal clinically important differences were considered in two contexts when interpreting the
35 diagnostic evidence in this guideline. When considering individual factors in isolation, it was
36 agreed in collaboration with the GDG that a positive likelihood ratio of 2 would constitute
37 significant diagnostic value. Therefore, when interpreting the diagnostic accuracy results for
38 single factors in isolation, something would only be considered to have diagnostic value if the
39 result was statistically significant at the 95% confidence level, and the point estimate of the
40 positive likelihood ratio was greater than 2 (or equivalently, value at ruling out the disease if
41 the negative likelihood ratio was less than 0.5).

When considering models containing multiple factors, the individual predictive value of each factor in isolation was considered not be a meaningful measure, as if the joint effect of a number of factors, which may have only limited diagnostic value individually, is to create an overall algorithm which is highly predictive, considering the diagnostic utility of the individual factors is no longer relevant. Further, there may well be correlations/interactions between factors which means the overall diagnostic value of two factors may be considerably different (in either direction) from the value one would predict be simply assuming independence of the individual elements. Therefore, MCIDs were not considered as part of the process of assessing algorithms containing multiple factors.

4.2.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

4.2.3 Evidence statements

4.2.3.1 People with predominantly axial symptoms

CONFIRMING that a person with predominantly axial symptoms has spondyloarthritis

On their own, the following findings increase the probability that a person with predominantly axial symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **large** increase in probability:

- Low-quality evidence:
  - HLA-B27 positivity (13 studies; total n=4,645)

EXCLUDING the possibility that a person with predominantly axial symptoms has spondyloarthritis

On their own, the following findings decrease the probability that a person with predominantly axial symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** decrease in probability:

- Very low-quality evidence:
  - HLA-B27 negativity (13 studies; total n=4,645)

Findings that DO NOT CLEARLY CONFIRM OR EXCLUDE the possibility that a person with predominantly axial symptoms has spondyloarthritis

On their own, the following findings do not clearly alter the probability that a person with predominantly axial symptoms has spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in probability:

- Low-quality evidence:
  - Normal CRP (5 studies; total n=2,389)
  - Normal ESR (1 study; total n=92)
- Very low-quality evidence:
  - Elevated CRP (5 studies; total n=2,389)
  - Elevated ESR (1 study; total n=92)

4.2.3.2 People with predominantly peripheral symptoms

CONFIRMING that a person with predominantly peripheral symptoms has spondyloarthritis

On their own, the following findings increase the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase or large increase in risk:

- Very low-quality evidence:
  - HLA-B27 positivity (7 studies; total n=1,005).

On their own, the following findings increase the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a **slight** degree:

- Moderate-quality evidence:
  - Elevated CRP (2 studies; total n=412).

EXCLUDING the possibility that a person with predominantly peripheral symptoms has spondyloarthritis

On their own, the following findings decrease the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in risk:

- Moderate-quality evidence:
  - Normal CRP (2 studies; total n=412).
- Very low-quality evidence:
  - HLA-B27 negativity (7 studies; total n=1,005).

4.2.3.3 All spondyloarthritis

CONFIRMING that a person has spondyloarthritis

On their own, the following findings increase the probability that a person has spondyloarthritis to a **moderate** degree:

- Very low-quality evidence:
  - HLA-B27 positivity (30 studies; total n=8,125).

On their own, the following findings increase the probability that a person has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** or **large** increase in risk:

- Very low-quality evidence:
  - Elevated ESR (2 studies; total n=867).

On their own, the following findings increase the probability that a person has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** increase in risk:

- Very low-quality evidence:
  - Elevated CRP (8 studies; total n=3,576).

EXCLUDING the possibility that a person has spondyloarthritis

On their own, the following findings decrease the probability that a person has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** decrease in risk:

- Very low-quality evidence:
  - HLA-B27 negativity (30 studies; total n=8,125).

On their own, the following findings decrease the probability that a person has spondyloarthritis to a **slight** degree:

- High-quality evidence:
  - Normal ESR (2 studies; total n=867).

Findings that DO NOT CLEARLY CONFIRM OR EXCLUDE the possibility that a person has spondyloarthritis

On their own, the following findings do not clearly alter the probability that a person has spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in probability:

- Low-quality evidence:
  - Normal CRP (8 studies; total n=3,576).

4.2.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that for all diagnostic questions, the most useful measures of diagnostic accuracy were positive and negative likelihood ratios. A positive likelihood ratio greater than 2 and a negative likelihood ratio less than 0.5 were considered to be clinically important results.
Trade-off between benefits and harms	The main benefit to carrying out these tests was to provide additional information to support a referral decision or a diagnosis of spondyloarthritis. The primary risks of these tests was agreed to be misinterpretation of the results, particularly leading to spondyloarthritis being ruled out inappropriately. This is particularly the case with HLA-B27 where some practitioners may misconstrue it to be an essential diagnostic criterion for spondyloarthritis.
Economic considerations	No economic evidence was identified for this review question. A referral strategy limited to HLA-B27 results alone was considered as part of the original consideration of referral strategies for axial disease undertaken for this guideline – see 4.1.4. The use of HLA-B27 (when not already performed), ESR and CRP assays was also part of the diagnostic work-up simulated by the original referral model (see Appendix H).
Quality of evidence	HLA-B27: Thirty papers were included which presented evidence on HLA-B27. The evidence was judged to be of low to very low quality. Thirteen studies presented data in axial populations, 7 in peripheral populations, and 10 in mixed axial and peripheral populations. The GDG discussed the evidence for axial disease and noted that the positive likelihood ratio was above the threshold for clinical significance, and the negative likelihood ratio was on the border of it. It was noted that the results for peripheral disease were less strong, though still statistically significant. The results for mixed populations of both axial and peripheral disease had a clinically significant result for the positive likelihood ratio but not for the negative likelihood ratio.

The GDG noted that the finding that HLA-B27 was positively associated with a diagnosis of peripheral disease represented a different perspective to the conventional view that HLA-B27 positivity is primarily of use when investigating axial spondyloarthritis.

The GDG also noted that 1 study (van Hoeven 2014) had far lower rates of HLA-B27 in both groups than they would expect. The authors did not identify a reason for this.

Erythrocyte sedimentation rate (ESR):
Two included studies presented evidence on ESR. The evidence was judged to be of low to very low quality in 1 study in axial disease, where the positive and negative likelihood ratios were not statistically significant in both cases. In the other study, where the population was a mixture of people with axial and peripheral disease, high quality evidence showed a weak clinically important positive likelihood ration, but a non-clinically important negative likelihood ratio. When pooled, neither the negative nor positive likelihood ratios were clinically important values.

C-reactive protein (CRP):
Eight included studies presented evidence on CRP. The evidence was judged to be of low to very low quality in axial populations, moderate quality in peripheral populations and high quality in mixed populations. The GDG noted that overall the evidence of diagnostic utility of CRP was weak, with no utility found in axial or mixed populations and slight utility in peripheral populations.

The GDG noted that while ESR and CRP were useful at other stages of care (e.g. evaluating treatment responses) they have less of a role to play in diagnosis. Nonetheless a positive result may provide additional supporting evidence during diagnostic investigations.

Overall it was agreed that a negative result from any of the above tests was insufficient evidence to rule out spondyloarthritis. The GDG agreed that at present there are cases of people with negative HLA-B27 results having a diagnosis of spondyloarthritis ruled out, even in the presence of symptoms which would suggest suspicion of spondyloarthritis. It was agreed that none of these tests should be used in isolation to make a referral or diagnostic decision, but rather be considered alongside other information about presenting signs, symptoms, risk factors and results of investigations.

Other considerations	Based on the evidence available, the GDG recommended that HLA-B27, and ESR/CRP testing may be part of the diagnostic work up for spondyloarthritis, but a negative result for this or ESR/CRP testing should not be used to rule out spondyloarthritis. It was noted that prevalence of HLA-B27 positivity varies by ethnic group. Presently the HLA-B27 test is not universally approved for all GPs to be able to order; practice varies between Clinical Commissioning Groups. The GDG were not able to make recommendations around whether the test should be carried out in primary versus secondary care (or either) on the basis of the evidence presented.

1

4.2.5 Recommendations

7. Blood tests for spondyloarthritis

7.1. Do not rule out a diagnosis of spondyloarthritis solely on the basis of a negative HLA-B27 result.

7.2. Do not rule out a diagnosis of spondyloarthritis even if a person's C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) are normal.

4.3 Imaging for diagnosis of spondyloarthritis

Review Question 10

What is the diagnostic utility of imaging (alone or in sequence) for investigating suspected spondyloarthritis?

4.3.1 Evidence review

The aim of this review was to assess the utility of imaging techniques as part of the recognition or diagnostic process for people with suspected spondyloarthritis.

The review focused on identifying studies that fulfilled the conditions specified in Table 15.

Table 15: PICO table for question 10: Imaging for diagnosis

Population	Patients with suspected spondyloarthritis
Interventions	- MRI - X-ray - Ultrasound - Isotope bone scan - PET CT - PET MRI - Sequential combinations of the above
Comparators	Clinical opinion of spondyloarthritis was considered the preferred reference standard, with diagnosis using any specified criteria as the next preference
Outcomes	- Sensitivity - Specificity - Positive likelihood ratio - Negative likelihood ratio - Positive predictive value - Negative predictive value - Diagnostic odds ratio

For full details of the review protocol please see Appendix C.

Cross-sectional studies and cohort studies were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well.

Systematic searches for this question identified 7,177 references, which were screened on their titles and abstracts. Between papers identified from this search and papers identified from other diagnostic searches in this guideline which contained relevant data, a total of 46 studies were retrieved for full text review. Thirty three of these studies were excluded as they did not meet the eligibility criteria such as inappropriate study design (e.g. case reports, case series with fewer than ten cases) or non-primary studies (e.g. systematic review, editorial). Detailed lists of excluded studies and reasons for their exclusion are provided in Appendix F.

4.3.1.1 Description of included studies

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G. All of the included studies were cross sectional.

4.3.1.1.1 *X-ray*

Nine studies were identified which contained relevant information on the diagnostic value of sacroiliitis detected on x-ray (3 in axial spondyloarthritis, 5 in peripheral spondyloarthritis and 1 in a mixed population containing people with both axial and peripheral spondyloarthritis).

One study was identified which contained relevant information on the diagnostic value of finger or toe pathology detected on x-ray (peripheral spondyloarthritis).

Two studies were identified which contained relevant information on the diagnostic value of enthesitis detected on x-ray (1 in peripheral spondyloarthritis and 1 in a mixed population containing people with both axial and peripheral spondyloarthritis).

4.3.1.1.2 *MRI*

Five studies were identified which contained relevant information on the diagnostic value of sacroiliitis detected on MRI (3 in axial spondyloarthritis, 1 in peripheral spondyloarthritis and 1 in a mixed population containing people with both axial and peripheral spondyloarthritis).

One study was identified which contained relevant information on the diagnostic value of spinal features detected on MRI (axial spondyloarthritis).

One study was identified which contained relevant information on the diagnostic value enthesitis detected on MRI (mixed population containing people with both axial and peripheral spondyloarthritis).

4.3.1.1.3 *Ultrasound*

One study was identified which contained relevant information on the diagnostic value of finger or toe pathology detected on ultrasound (peripheral spondyloarthritis).

One study was identified which contained relevant information on the diagnostic value of finger or toe pathology detected on power Doppler ultrasound (peripheral spondyloarthritis).

One study was identified which contained relevant information on the diagnostic value of enthesitis detected on power Doppler ultrasound (mixed population containing people with both axial and peripheral spondyloarthritis).

4.3.1.1.4 *Scintigraphy*

One study was identified which contained relevant information on the diagnostic value of sacroiliitis detected on scintigraphy (axial spondyloarthritis).

4.3.1.2 Variations from protocol

A specific search was conducted for this question, to identify studies which sought to evaluate the diagnostic utility different imaging modalities. However, in the course of conducting other diagnostic utility questions for this guideline, further data was identified which was contained incidentally in other studies. Across all of the searches conducted for any of the diagnostic questions, any relevant data was extracted, if the study met the eligibility criteria, regardless of whether it was identified in the specific search for this study.

After drafting the original protocol, the GDG opted to narrow the range of eligible imaging modalities further. PET scanning was excluded due to its use primarily being restricted to research in this population, and CT scanning was excluded as it is only used in a small number of cases where MRI is contra-indicated.

4.3.1.3 Minimal clinically important differences

Minimal clinically important differences were considered in two contexts when interpreting the diagnostic evidence in this guideline. When considering individual factors in isolation, it was agreed in collaboration with the GDG that a positive likelihood ratio of 2 would constitute significant diagnostic value. Therefore, when interpreting the diagnostic accuracy results for single factors in isolation, something would only be considered to have diagnostic value if the result was statistically significant at the 95% confidence level, and the point estimate of the positive likelihood ratio was greater than 2 (or equivalently, value at ruling out the disease if the negative likelihood ratio was less than 0.5).

When considering models containing multiple factors, the individual predictive value of each factor in isolation was considered not be a meaningful measure, as if the joint effect of a number of factors, which may have only limited diagnostic value individually, is to create an overall algorithm which is highly predictive, considering the diagnostic utility of the individual factors is no longer relevant. Further, there may well be correlations/interactions between factors which means the overall diagnostic value of two factors may be considerably different (in either direction) from the value one would predict be simply assuming independence of the individual elements. Therefore, MCIDs were not considered as part of the process of assessing algorithms containing multiple factors.

4.3.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question.

Health economic modelling was not directly prioritised for this review question; however, evidence from some of the identified research was considered as part of the modelling of potential referral strategies for axial spondyloarthritis (see 4.1.2.2).

4.3.3 Evidence statements

4.3.3.1 People with predominantly axial symptoms

CONFIRMING that a person with predominantly axial symptoms has spondyloarthritis

On their own, the following findings increase the probability that a person with predominantly axial symptoms has spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **large** increase in risk:

- Moderate-quality evidence:
  - Sacroiliitis on MRI (3 studies; total n=1,550).
  - Sacroiliitis on X-ray (4 studies; total n=1,762).

On their own, the following findings increase the probability that a person with predominantly axial symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase in risk:

- Low-quality evidence:
  - Spinal pathology on MRI (1 study; total n=708).

On their own, the following findings increase the probability that a person with predominantly axial symptoms has spondyloarthritis to a **slight** degree:

- Moderate-quality evidence:
  - Sacroiliitis on scintigraphy (1 study; total n=194).

EXCLUDING the possibility that a person with predominantly axial symptoms has spondyloarthritis

On their own, the following findings decrease the probability that a person with predominantly axial symptoms has spondyloarthritis to a **slight** degree:

- Low-quality evidence:
  - Absence of sacroiliitis on X-ray (4 studies; total n=1,762).
- Moderate-quality evidence:
  - Absence of sacroiliitis on MRI (3 studies; total n=1,550).
  - Absence of spinal pathology on MRI (1 study; total n=708).

On their own, the following findings decrease the probability that a person with predominantly axial symptoms has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in risk:

- Low-quality evidence:
  - Absence of sacroiliitis on scintigraphy (1 study; total n=194).

4.3.3.2 People with predominantly peripheral symptoms

CONFIRMING that a person with predominantly peripheral symptoms has spondyloarthritis

On their own, the following findings increase the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **moderate** or **large** increase in risk:

- High-quality evidence:
  - Finger or toe pathology on ultrasound (1 study; total n=52).

On their own, the following findings increase the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **large**; however, at a 95% confidence level, data are also consistent with a **slight**, **moderate** or **very large** increase in risk:

- Very low-quality evidence:
  - Sacroiliitis on X-ray (4 studies; total n=539).

On their own, the following findings increase the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase in risk:

- Moderate-quality evidence:
  - Finger or toe pathology on power Doppler ultrasound (1 study; total n=52).

EXCLUDING the possibility that a person with predominantly peripheral symptoms has spondyloarthritis

On their own, the following findings decrease the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **moderate** or **large** decrease in risk:

- High-quality evidence:
  - Absence of finger or toe pathology on ultrasound (1 study; total n=52).

On their own, the following findings decrease the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** or **large** decrease in risk:

- Moderate-quality evidence:
    - o Absence of finger or toe pathology on power Doppler ultrasound (1 study; total n=52).

On their own, the following findings decrease the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a **slight** degree:

- High-quality evidence:
    - o Absence of finger or toe pathology on X-ray (1 study; total n=52).
- Low-quality evidence:
    - o Absence of sacroiliitis on X-ray (4 studies; total n=539).

On their own, the following findings decrease the probability that a person with predominantly peripheral symptoms has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in risk:

- Moderate-quality evidence:
    - o Absence of enthesitis on X-ray (1 study; total n=81).
    - o Absence of sacroiliitis on MRI (1 study; total n=60).

Findings that DO NOT CLEARLY CONFIRM OR EXCLUDE the possibility that a person with predominantly peripheral symptoms has spondyloarthritis

On their own, the following findings do not clearly alter the probability that a person with predominantly peripheral symptoms has spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in probability:

- Moderate-quality evidence:
    - o Enthesitis on X-ray (1 study; total n=81).
    - o Finger or toe pathology on X-ray (1 study; total n=52).
    - o Sacroiliitis on MRI (1 study; total n=60).

4.3.3.3 All spondyloarthritis

CONFIRMING that a person has spondyloarthritis

On their own, the following findings increase the probability that a person has spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **large** increase in risk:

- Moderate-quality evidence:
    - o Sacroiliitis on MRI (5 studies; total n=1,683).
    - o Sacroiliitis on X-ray (9 studies; total n=3,076).

EXCLUDING the possibility that a person has spondyloarthritis

On their own, the following findings decrease the probability that a person has spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in risk:

- Low-quality evidence:
    - o Absence of enthesitis on X-ray (2 studies; total n=114).

On their own, the following findings decrease the probability that a person has spondyloarthritis to a **slight** degree:

- Low-quality evidence:
  o Absence of sacroiliitis on X-ray (9 studies; total n=3,076).
- Moderate-quality evidence:
  o Absence of sacroiliitis on MRI (5 studies; total n=1,683).

Findings that DO NOT CLEARLY CONFIRM OR EXCLUDE the possibility that a person has spondyloarthritis

On their own, the following findings do not clearly alter the probability that a person has spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in probability:

- Very low-quality evidence:
  o Enthesitis on X-ray (2 studies; total n=114).

4.3.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that for all diagnostic questions, the most useful measures of diagnostic accuracy were positive and negative likelihood ratios. A positive likelihood ratio greater than 2 and a negative likelihood ratio less than 0.5 were considered to be clinically important results.
	The GDG had previously received advice from a co-opted radiologist regarding which papers were reporting imaging investigations relevant to clinical practice and which were research-only measures, leading to an agreement as to what papers would be excluded from consideration. For this reason, PET scanning was not used. CT-scanning was also excluded from the evidence as the GDG stated it is only used in small subset of people for whom MRI is contraindicated.
Trade-off between benefits and harms	There is a risk of delay to diagnosis if the wrong kind of imaging is requested, or it is not interpreted by a specialist with knowledge of spondyloarthritis. This may lead to an avoidable repeat of imaging being required.
	The GDG discussed what should happen in the event of a negative result on imaging, to prevent people from prematurely having spondyloarthritis ruled out if other symptoms are present. It was acknowledged that, in some people, a repeat imaging investigation may be needed further down the line when detectable features may have developed.
	The GDG noted that radiation exposure people receive when undergoing diagnostic investigations with X-ray may pose a health risk, though sacroiliac joint imaging is a lower dose than some imaging. It is nonetheless considered to be an appropriate diagnostic tool, though in the absence of radiographic findings, careful consideration of the risks of X-ray should be made before requesting a repeat.
	While MRI is generally considered a more sensitive method of detecting sacroiliitis, access to timely and appropriate MRI needs to be considered, and costs also factored into the decision to request an MRI. It was also noted by the GDG that people sometimes are concerned about undergoing an MRI scan and consequently either do not attend, or find the experience too difficult to tolerate.
Economic considerations	The costs of diagnostic imaging were included in the original model investigating potential referral strategies, and their benefits were inherent in the accuracy of those strategies (as judged against a final

	diagnosis). Therefore, it can be inferred that, although different approaches to diagnostic imaging were not a decision-point in the model, any referral strategy that provides good value for money must be making reasonable use of these resources.
Quality of evidence	**Axial imaging** The GDG discussed the presented evidence and noted the strong positive likelihood ratios for use of both X-ray and MRI of the sacroiliac joints as part of the diagnostic work up for suspected axial spondyloarthritis, and made recommendations for the use of each. It was noted that evidence on spinal MRI was also available, but this evidence did not show a clinically significant result and therefore a recommendation was not made on spinal imaging. Low to moderate quality evidence found statistically significant but not clinically important positive and negative likelihood ratios for the detection of sacroiliitis using scintigraphy. The GDG noted that the superior performance of both X-ray and MRI meant that they were more appropriate for first and second line investigation respectively. In the absence of findings using both X-ray and MRI, where axial disease is still suspected, the GDG agreed it was more appropriate to wait and perform a further X-ray/MRI at a later date rather than use a third imaging approach, as the initial absence of findings may reflect the disease stage more than the mode of imaging used. The GDG noted that some of the presented evidence suggested that axial signs on imaging can be detected in people presenting with suspected peripheral disease, even in the absence of reported back pain. Identifying axial involvement is important when selecting what treatments to offer. They also discussed that, in their experience, people with longstanding spinal stiffness may no longer be aware of what normal back flexion feels like, and people with a lot of fusion might not experience back pain any more. The need for the correct type of MRI was discussed, particularly in sacroiliac joint imaging, and the GDG opted to be explicit about the type in the recommendations. They agreed that MRI scans need to be interpreted according to Assessment of Spondyloarthritis (ASAS) MRI recommendations. **Peripheral imaging** The GDG agreed that the evidence suggested that finger and toe ultrasound was useful in people presenting with peripheral symptoms, but there was neither evidence to suggest this approach should be extended to all peripheral joints, nor would it be practical to recommend conducting ultrasound investigation on all joints in people presenting with suspected peripheral disease. Therefore, the GDG agreed that imaging of other sites should be restricted to the hands and feet (which can be done as part of the same scan as the fingers and toes) and symptomatic sites (i.e. where there is good reason to suppose inflammation is present). They agreed it would not be appropriate (both in terms of costs and patient burden) to recommend routine imaging of non-symptomatic sites.
Other considerations	<u>Axial imaging</u> The GDG made recommendations regarding the use of X-ray and MRI in diagnosis of axial spondyloarthritis based on the presented evidence. They additionally drew on their professional experience and that of co-opted specialist radiologists to decide the sequence of imaging and what action should be taken in the event of a negative finding on imaging. The GDG felt it was appropriate to recommend X-ray as first line imaging as it is more accessible, and is sufficient to diagnoses

ankylosing spondylitis. They agreed that only when sacroiliitis was not detectable on X-ray should people who are still considered to potentially have spondyloarthritis receive and MRI.

The GDG discussed whether the type of MRI scan requested had an influence on the likelihood of sacroiliitis being detected. Though no evidence comparing different MRI requests was presented, the GDG were confident in making a recommendation on the basis of their clinical experience, particularly as guided by two radiologists.

It was agreed that the modified New York radiographic criteria should be used to interpret X-ray findings, in line with current practice. On the advice of the co-opted experts, the GDG agreed that ASAS-MRI criteria should be used to interpret MRI findings. These criteria additionally advise use of HLA-B27 testing in the event of a negative finding, if this has not already been investigated.

The GDG decided that scintigraphy was relatively poor performing and that, if sacroiliitis was not detected on X-ray or MRI but spondyloarthritis was still suspected, it would be more beneficial to wait and repeat those forms of imaging rather than use scintigraphy.

Peripheral imaging
The GDG also made recommendations regarding the use of both peripheral joint and spinal imaging in diagnosis of peripheral spondyloarthritis, based on the evidence presented. The sequence of imaging was based on the experience of the GDG and co-opted radiologists.
The recommendation to offer imaging of hands and feet to people presenting with peripheral symptoms is in line with current practice.

4.3.5   Recommendations

8.   Imaging for diagnosis of peripheral spondyloarthritis

8.1.   Offer plain film X-ray of the sacroiliac joints for people with suspected axial spondyloarthritis.

8.2.   Diagnose axial spondyloarthritis if the plain film X-ray shows sacroiliitis meeting the modified New York radiological criteria (bilateral grade 2-4 or unilateral grade 3-4 sacroiliitis).

8.3.   If the plain film X-ray does not show sacroiliitis meeting modified New York radiological criteria (bilateral grade 2-4 or unilateral grade 3-4 sacroiliitis), request unenhanced MRI using an inflammatory back pain protocol.

8.4.   Radiologists receiving a request for an inflammatory back pain MRI should perform short T1 inversion recovery (STIR), T1 (both views), cervical, thoracic and lumbar (whole spine, sagittal view), and sacroiliac joints (coronal view).

8.5.   Use the ASAS/OMERACT MRI criteria to interpret the MRI.

8.6.   Diagnose axial spondyloarthritis if the MRI meets the ASAS/OMERACT MRI criteria.

8.7.   If the MRI does not meet the ASAS/OMERACT MRI criteria:

- do not exclude the possibility of axial spondyloarthritis

- consider specialist musculoskeletal radiology review if there is disparity between the clinical suspicion and imaging findings, particularly in people with an immature skeleton.

8.8. **If interpretation of MRI findings remains negative, offer an HLA-B27 test if it has not already been done.**

8.9. **If the HLA-B27 test is positive, base the diagnosis on clinical features, for example, using the clinical 'arm' of the ASAS axial classification criteria.**

8.10. **If a diagnosis of spondyloarthritis cannot be made and clinical suspicion remains high, consider a follow-up MRI.**

8.11. **Do not offer scintigraphy for people with suspected axial spondyloarthritis.**

9. **Imaging for diagnosis of peripheral spondyloarthritis**

9.1. **Offer plain film X-ray of symptomatic hands and feet for people with suspected peripheral spondyloarthritis in these areas.**

9.2. **If a diagnosis cannot be made from the plain film X-ray, consider ultrasound:**

- of the hands and feet to assess for joint involvement
- of suspected enthesitis sites.

9.3. **Consider plain film X-rays, ultrasound and/or MRI of other peripheral symptomatic sites.**

9.4. **If a diagnosis of peripheral spondyloarthritis is confirmed:**

- offer plain film X-ray of the sacroiliac joint to assess for axial involvement, even if the person does not have any symptoms
- only consider MRI of the sacroiliac joint if the result is likely to change management.

9.5. **Interpret a positive HLA-B27 result as increasing the likelihood of peripheral spondyloarthritis**

4.4 Information gathering to improve early diagnosis

Review Question 5

- What is the usefulness of information gathering (for example family history, self-report questionnaires, and screening criteria) in improving early diagnosis of spondyloarthritis?

4.4.1 Evidence review

The aim of this review question was to ascertain the utility of routinely collecting information prior to making a diagnosis.

The review focused on identifying studies that fulfilled the conditions specified in Table 16.

Table 16: PICO table information gathering

Population	Patients with a suspected or confirmed diagnosis of spondyloarthritis, healthcare professionals
Interventions	Barriers such as: • Lack of patient awareness leading to delayed diagnosis • Patients deterred by lack of diagnosis at earlier consultation • Lack of health-care professional awareness of chronic inflammatory conditions • Lack of health-care professional awareness of complications/comorbid manifestations of pre-existing inflammatory conditions • High consultation rate of lower back pain (mostly mechanical) • Lack of cross referrals in secondary care between relevant specialities • Over-specialism within rheumatology leading to consultations where relevant comorbidities are not assessed. • Lack of multidisciplinary team assessment • Lack of access from GPs to (i) HLA-B27 testing (ii) appropriate MRI equipment or protocol • Patient gender (under-diagnosis in women) • Lack of a biological marker in spondyloarthritis
Comparators	Prompt diagnosis of spondyloarthritis
Outcomes	Time to appointment, number of contacts with health care professionals, health related quality of life, resource use and costs, patient satisfaction, disease burden reduced from both spondyloarthritis and associated conditions, service delivery/organisation

For full details of the review protocol please see Appendix C.

Potential study designs that may have helped to answer this question and were in the review protocol included observational intervention study designs.

A systematic search identified 7,813 references. The references were screened on their titles and abstracts and 11 references were ordered for full text; none of these met the inclusion criteria for this review question. A list of excluded studies and the reasons for their exclusion is available in Appendix F.

4.4.1.1 Description of included studies

No studies met the inclusion criteria for this review.

4.4.2 Health economics evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A

1 total of 9,970 references was retrieved, of which none were retained for this review question.
2 Health economic modelling was not prioritised for this review question.

3 **4.4.3 Evidence statements**

4 No evidence was identified for this review question.

5 **4.4.4 Evidence to recommendations**

Relative value of different outcomes	In the absence of any published evidence, the GDG used their experience and expertise to discuss the information which should be collected from people presenting with suspected spondyloarthritis. It noted that, in addition to personal medical history of established comorbidities and associated conditions (uveitis, psoriasis, IBD), it is important to gather information about family history of these conditions.
	The importance of skilled history taking was discussed; it was noted that asking people about a personal or family history of psoriasis, for example, may be less useful than enquiring about 'skin problems' with appropriate supplementary questions. The GDG agreed that, when asking people with suspected SpA about their symptoms, it is important to prompt them to think about past symptoms as well as current ones (e.g. people may have had psoriasis in the past but not mention it as it has not occurred recently).
	Key features for distinguishing between axial SpA and mechanical back pain were noted by the GDG to include the nature of back pain at rest or under exercise, mode of onset, and alternation of site of pain.
	It was noted that, although presentation following trauma is usually indicative of mechanical back pain (whereas inflammatory back pain typically has an insidious onset), some people with inflammatory back pain only seek medical advice when their underlying condition is exacerbated by an accident or other sudden trigger. The GDG therefore agreed on the importance of eliciting a full history of relevant symptoms, rather than relying on the person's current description only.
Trade-off between benefits and harms	No substantial harms were identified in asking people with suspected spondyloarthritis about their condition.
	It was noted that some people with insidious onset of back pain adapt to the pain and any limitations to their activity such that they may under-report their symptoms, as they may have come to perceive their symptoms as 'normal'.
	The GDG noted that there are several validated instruments available for diagnosing inflammatory back pain (IBP), but that these may still fail to detect all people with IBP, so should not be used to make a decision to refer or diagnose a patient without being supported by other pieces of information or results of investigations.
Economic considerations	No health economic evidence was identified for this question.
Quality of evidence	No clinical evidence was identified on the usefulness of information gathering. The GDG did not consider there to be additional value in having a separate recommendation on information gathering, when many of the relevant features are included in the recommendations on signs, symptoms and risk factors of spondyloarthritis.
Other considerations	Not applicable

6 **4.4.5 Recommendations**

7 No recommendations or research recommendations were made.

4.5 Diagnostic risk scores and models

Review Question 4

- What is the diagnostic utility of a risk assessment score for identifying spondyloarthritis?

4.5.1 Evidence review

The aim of this review was to assess the utility of risk assessment scores, models or tools as part of the recognition or diagnostic process for people with suspected spondyloarthritis.

The review focused on identifying studies that fulfilled the conditions specified in Table 17.

Table 17: PICO table for question 4: Diagnostic risk scores

Population	Patients with suspected spondyloarthritis

Interventions	• Any risk assessment score/rule/model presenting at least two characteristics in combination
	• Clinical diagnostic scores/tools such as:
	  o Axial
	    - Modified New York criteria
	    - Bennett's criteria etc.
	    - ASAS criteria (axial)
	    - AMOR criteria
	    - Calin criteria
	    - European Spondyloarthropathy Study Group criteria
	  o Peripheral
	    - CASPAR criteria
	    - ASAS criteria (peripheral)
	    - Modified McGonagle criteria
	  o Both
	    - Moll and Wright
	    - Vasey and Espinoza

Comparators	Clinical opinion of spondyloarthritis
Outcomes	• Sensitivity • Specificity • Positive likelihood ratio • Negative likelihood ratio • Positive predictive value • Negative predictive value • Diagnostic odds ratio

For full details of the review protocol please see Appendix C.

Cross-sectional studies and cohort studies were considered to be the highest quality evidence available to answer this question and are graded as high in a modified GRADE framework if conducted and reported well.

Systematic searches for this question identified 7,953 references, which were screened on their titles and abstracts. Between papers identified from this search and papers identified from other diagnostic searches in this guideline which contained relevant data, a total of 162 studies were retrieved for full text review. 152 of these studies were excluded as they did not meet the eligibility criteria such as inappropriate study design (e.g. case reports, case series with fewer than ten cases) or non-primary studies (e.g. editorials). Detailed lists of excluded studies and reasons for their exclusion are provided in Appendix F Ten cross sectional studies were included in the review, with descriptions of the criteria evaluated given below.

4.5.1.1 Description of included studies

Six studies were identified which contained relevant information on the diagnostic accuracy of the AMOR criteria (2 in axial spondyloarthritis, 1 in peripheral spondyloarthritis and 3 in a mixed population containing people with both axial and peripheral spondyloarthritis).

Three studies were identified which contained relevant information on the diagnostic accuracy of the ASAS axial criteria (2 in axial spondyloarthritis and 1 in a mixed population containing people with both axial and peripheral spondyloarthritis).

Three studies were identified which contained relevant information on the diagnostic accuracy of the Berlin criteria (2 in axial spondyloarthritis and 1 in a mixed population containing people with both axial and peripheral spondyloarthritis).

Six studies were identified which contained relevant information on the diagnostic accuracy of the ESSG criteria (2 in axial spondyloarthritis, 1 in peripheral spondyloarthritis and 1 in a mixed population containing people with both axial and peripheral spondyloarthritis).

Two studies were identified which contained relevant information on the diagnostic accuracy of the New York and modified New York criteria (both in axial spondyloarthritis).

One study was identified which contained relevant information on the diagnostic accuracy of the Rome criteria (axial spondyloarthritis).

One study was identified which contained relevant information on the diagnostic accuracy of the ASAS peripheral criteria (peripheral spondyloarthritis).

One study was identified which contained relevant information on the diagnostic accuracy of the French Society for Rheumatology criteria (peripheral spondyloarthritis).

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G.

4.5.1.2 Variations from protocol

A specific search was conducted for these questions, to identify studies which sought to evaluate the diagnostic utility of diagnostic tools and models. However, in the course of conducting other diagnostic utility questions for this guideline, further data were identified which was contained incidentally in other studies. Across all of the searches conducted for any of the diagnostic questions, any relevant data were extracted, if the study met the eligibility criteria, regardless of whether it was identified in the specific search for this study.

4.5.1.3 Minimal clinically important differences

Minimal clinically important differences were considered in two contexts when interpreting the diagnostic evidence in this guideline. When considering individual factors in isolation, it was agreed in collaboration with the GDG that a positive likelihood ratio of 2 would constitute significant diagnostic value. Therefore, when interpreting the diagnostic accuracy results for single factors in isolation, something would only be considered to have diagnostic value if the result was statistically significant at the 95% confidence level, and the point estimate of the positive likelihood ratio was greater than 2 (or equivalently, value at ruling out the disease if the negative likelihood ratio was less than 0.5).

When considering models containing multiple factors, the individual predictive value of each factor in isolation was considered not be a meaningful measure, as if the joint effect of a number of factors, which may have only limited diagnostic value individually, is to create an overall algorithm which is highly predictive, considering the diagnostic utility of the individual factors is no longer relevant. Further, there may well be correlations/interactions between factors which means the overall diagnostic value of two factors may be considerably different

1 (in either direction) from the value one would predict be simply assuming independence of
2 the individual elements. Therefore, MCIDs were not considered as part of the process of
3 assessing algorithms containing multiple factors.

4 ### 4.5.2 Health economic evidence

5 A literature search was conducted jointly for all review questions in this guideline by applying
6 standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
7 total of 9,970 references was retrieved, of which none were retained for this review question.

8 Health economic modelling was not directly prioritised for this review question; however,
9 evidence from some of the identified research was considered as part of the modelling of
10 potential referral strategies for axial spondyloarthritis (see 4.1.2.2).

11 ### 4.5.3 Evidence statements

12 **CONFIRMING that a person with predominantly axial symptoms has spondyloarthritis**

13 If a person with predominantly axial symptoms **fulfils** the following criteria, it increases the
14 probability that they have spondyloarthritis to a **very large** degree:

15 - High-quality evidence:
16   - ASAS axial criteria (imaging arm only) (1 study; total n=649).
17 - Moderate-quality evidence:
18   - Rome radiographic criteria (1 study; total n=212).

19 If a person with predominantly axial symptoms **fulfils** the following criteria, it increases the
20 probability that they have spondyloarthritis to a degree that is most likely to be **very large**;
21 however, at a 95% confidence level, data are also consistent with a **large** increase in
22 probability:

23 - Moderate-quality evidence:
24   - Original New York criteria (1 study; total n=212).

25 If a person with predominantly axial symptoms **fulfils** the following criteria, it increases the
26 probability that they have spondyloarthritis to a **moderate** degree:

27 - High-quality evidence:
28   - Berlin criteria (2 studies; total n=842).
29   - Berlin criteria (first van den Berg modification) (2 studies; total n=842).
30   - Berlin criteria (second van den Berg modification) (2 studies; total n=842).

31 If a person with predominantly axial symptoms **fulfils** the following criteria, it increases the
32 probability that they have spondyloarthritis to a degree that is most likely to be **moderate**;
33 however, at a 95% confidence level, data are also consistent with a **slight** increase in
34 probability:

35 - Low-quality evidence:
36   - Rome clinical criteria (1 study; total n=212).

37 **EXCLUDING the possibility that a person with predominantly axial symptoms has**
38 **spondyloarthritis**

39 If a person with predominantly axial symptoms **does not fulfil** the following criteria, it
40 decreases the probability that they have spondyloarthritis to a degree that is most likely to be
41 **large**; however, at a 95% confidence level, data are also consistent with a **moderate** or **very**
42 **large** decrease in probability:

43 - Moderate-quality evidence:

o Rome radiographic criteria (1 study; total n=212).

If a person with predominantly axial symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a **moderate** degree:

- High-quality evidence:
  - o ASAS axial criteria (imaging arm only) (1 study; total n=649).
  - o Berlin criteria (2 studies; total n=842).
  - o Berlin criteria (first van den Berg modification) (2 studies; total n=842).
  - o Berlin criteria (second van den Berg modification) (2 studies; total n=842).
  - o ESSG criteria (2 studies; total n=1,357).
- Moderate-quality evidence:
  - o Amor criteria (2 studies; total n=1,357).

If a person with predominantly axial symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **large** decrease in probability:

- Low-quality evidence:
  - o Modified Amor criteria (2 studies; total n=1,357).
  - o Modified ESSG criteria (2 studies; total n=1,357).

If a person with predominantly axial symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** or **large** decrease in probability:

- Low-quality evidence:
  - o Original New York criteria (1 study; total n=212).
- Very low-quality evidence:
  - o ASAS axial criteria (2 studies; total n=1,357).

Criteria that DO NOT CLEARLY CONFIRM OR EXCLUDE the possibility that a person with predominantly axial symptoms has spondyloarthritis

If a person with predominantly axial symptoms **fulfils** the following criteria, it does not clearly alter the probability that they have spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in probability:

- Very low-quality evidence:
  - o Amor criteria (2 studies; total n=1,357).
  - o Modified Amor criteria (2 studies; total n=1,357).
  - o ASAS axial criteria (2 studies; total n=1,357).
  - o ESSG criteria (2 studies; total n=1,357).
  - o Modified ESSG criteria (2 studies; total n=1,357).
  - o Modified New York criteria (2 studies; total n=920).

If a person with predominantly axial symptoms **does not fulfil** the following criteria, it does not clearly alter the probability that they have spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in probability:

- Moderate-quality evidence:
  - o Rome clinical criteria (1 study; total n=212).
- Very low-quality evidence:
  - o Modified New York criteria (2 studies; total n=920).

4.5.3.1 People with predominantly peripheral symptoms

CONFIRMING that a person with predominantly peripheral symptoms has spondyloarthritis

If a person with predominantly peripheral symptoms **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **large** increase in probability:

- Moderate-quality evidence:
  - French Society for Rheumatology criteria for reactive arthritis (1 study; total n=217).

If a person with predominantly peripheral symptoms **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a degree that is most likely to be **very large**; however, at a 95% confidence level, data are also consistent with a **moderate** or **large** increase in probability:

- High-quality evidence:
  - Amor criteria (1 study; total n=266).
  - Modified Amor criteria (1 study; total n=266).

If a person with predominantly peripheral symptoms **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **large** increase in probability:

- High-quality evidence:
  - ASAS peripheral criteria (1 study; total n=266).
  - Modified ESSG criteria (1 study; total n=266).

If a person with predominantly peripheral symptoms **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase in probability:

- Moderate-quality evidence:
  - ESSG criteria (1 study; total n=266).

EXCLUDING the possibility that a person with predominantly peripheral symptoms has spondyloarthritis

If a person with predominantly peripheral symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a **moderate** degree:

- High-quality evidence:
  - ASAS peripheral criteria (1 study; total n=266).

If a person with predominantly peripheral symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **large** decrease in probability:

- Moderate-quality evidence:
  - French Society for Rheumatology criteria for reactive arthritis (1 study; total n=217).

If a person with predominantly peripheral symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** decrease in probability:

- Moderate-quality evidence:
  - modified ESSG criteria (1 study; total n=266).

If a person with predominantly peripheral symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a **slight** degree:

- High-quality evidence:
  - Amor criteria (1 study; total n=266).
  - Modified Amor criteria (1 study; total n=266).

If a person with predominantly peripheral symptoms **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **slight**; however, at a 95% confidence level, data are also consistent with a **moderate** decrease in probability:

- Moderate-quality evidence:
  - ESSG criteria (1 study; total n=266).

4.5.3.2 All spondyloarthritis

CONFIRMING that a person has spondyloarthritis

If a person **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a **moderate** degree:

- High-quality evidence:
  - Berlin criteria (3 studies; total n=885).

If a person **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** increase in probability:

- Low-quality evidence:
  - ESSG criteria (6 studies; total n=2,530).
- Very low-quality evidence:
  - Modified ESSG criteria (3 studies; total n=1,623).

If a person **fulfils** the following criteria, it increases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** or **large** increase in probability:

- Low-quality evidence:
  - Amor criteria (6 studies; total n=2,530).
- Very low-quality evidence:
  - Modified Amor criteria (3 studies; total n=1,623).

EXCLUDING the possibility that a person has spondyloarthritis

If a person **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a **moderate** degree:

- High-quality evidence:
  - ASAS axial criteria (imaging arm only) (1 study; total n=649).
  - ASAS peripheral criteria (1 study; total n=266).
  - Berlin criteria (3 studies; total n=885).

If a person **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** decrease in probability:

- Low-quality evidence:
  - Amor criteria (6 studies; total n=2,530).
  - modified ESSG criteria (3 studies; total n=1,623).
- Moderate-quality evidence:
  - ESSG criteria (6 studies; total n=2,530).

If a person **does not fulfil** the following criteria, it decreases the probability that they have spondyloarthritis to a degree that is most likely to be **moderate**; however, at a 95% confidence level, data are also consistent with a **slight** or **large** decrease in probability:

- Low-quality evidence:
  - modified Amor criteria (3 studies; total n=1,623).
  - original New York criteria (1 study; total n=212).
- Very low-quality evidence:
  - ASAS axial criteria (3 studies; total n=1,400).

Criteria that DO NOT CLEARLY CONFIRM OR EXCLUDE the possibility that a person has spondyloarthritis

If a person **fulfils** the following criteria, it does not clearly alter the probability that they have spondyloarthritis; at a 95% confidence level, data are consistent with an increase or a decrease in probability:

- Very low-quality evidence:
     ASAS axial criteria (3 studies; total n=1,400).

4.5.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that for all diagnostic questions, the most useful measures of diagnostic accuracy were positive and negative likelihood ratios. A positive likelihood ratio greater than 2 and a negative likelihood ratio less than 0.5 were considered to be clinically important results.
Trade-off between benefits and harms	No direct harms were identified related to the use of scores and models as part of the diagnostic work-up on spondyloarthritis. However, some of the tests and imaging required by the scores may have harms associated with them (for example, exposure to radiation during X-ray, general small risks associated with blood tests). There is also a risk of incorrect diagnosis if a model is not sufficiently accurate, or is incorrectly applied. Potential benefits were identified as the possible increase in correct diagnoses.
Economic considerations	No health economic evidence was identified for this question. It was noted that, in the absence of a true reference standard against which to assess options, convincing evidence on the superiority of one diagnostic schema over another could only be provided by extended diagnose-and-treat comparative research, which is very unlikely to be commissioned. This was one of the reasons why no particular approach was recommended.
Quality of evidence	Evidence was presented on diagnostic tools, scores and models, including modified versions, as follows: 13 for axial spondyloarthritis, 6 for peripheral spondyloarthritis, and 9 for spondyloarthritis which did not specify whether the condition was axial or peripheral. There was considerable variation in the quality of the evidence, both between

	measures and between whether the same measure is being used to confirm or exclude spondyloarthritis (full quality ratings are given in the evidence statements above). The GDG noted that while many of the risk models and scores had statistically significant positive and/or negative likelihood ratios, not all of these measures were statistically significant. In common with other diagnostic reviews in this guideline, where a positive result is recorded by a score or a model, this may increase clinical suspicion or likelihood of diagnosis, but a negative result is insufficient evidence to rule out the condition. It was noted that the scores/models which required sacroiliitis findings on X-ray would, by definition, only diagnose ankylosing spondylitis and excluded non-radiographic axial spondyloarthritis, so these scores were not considered to be useful.
Other considerations	The GDG noted that where risk scores or models required imaging to be carried out, these tools would only be suitable for use in a secondary care setting, as there is varying provision in access to the correct type of imaging via primary care referrals. A similar variation of provision of access to HLA-B27 in primary care was also noted, though there are fewer barriers to changing this. The GDG noted that since none of the criteria were sufficiently accurate to be used to diagnose spondyloarthritis without specialist clinical input or specific clinical investigations (e.g. MRI), it would not be appropriate to use these tools for diagnostic purposes in primary care. The drafted recommendation reflected the evidence reviews that were presented. The decision not to advocate for the use of a particular tool was motivated by the need to not unduly restrict investigation into suspected spondyloarthritis at this stage of the care pathway, and the fact that no tool consistently outperformed the others in a way to justify prioritising it above the others.

4.5.5 Recommendations

10. In specialist settings, consider using validated spondyloarthritis criteria to diagnose spondyloarthritis. Examples of criteria include:

- Amor criteria
- ASAS axial or peripheral criteria
- Berlin criteria
- European Spondyloarthropathy Study Group (ESSG) criteria
- French Society of Rheumatology reactive arthritis criteria
- Rome criteria
- Modified New York criteria.

4.6 Microbiology testing in reactive arthritis

Review Question 11

- What is the diagnostic utility of testing for infection such as *Salmonella, Shigella, Yersinia, Campylobacter* and *Chlamydia* in cases of suspected reactive arthritis?

4.6.1 Evidence review

The aim of this review was to assess whether testing for specific pathogens in people with suspected reactive arthritis is sufficiently accurate to form a useful part of the diagnostic process.

The review focused on identifying studies that fulfilled the conditions specified in Table 18.

Table 18: PICO table for question 11: microbiology for reactive arthritis

Population	People (aged 16 years and over) with suspected reactive arthritis
Interventions	Specific testing/culture methods e.g. • Urine testing (*Chlamydia*) • Swabbing (*Chlamydia*) • Anal swabs (*Chlamydia*) • Blood cultures (all) • PCR (fragments of bacterial DNA) (all) • Faecal samples (GI infections)
Comparators	• Clinical diagnosis defined by specific criteria such as: o ASAS criteria (peripheral) • Clinical opinion
Outcomes	• Sensitivity • Specificity • Positive likelihood ratio • Negative likelihood ratio • Positive predictive value • Negative predictive value

For full details of the review protocol please see Appendix C.

Cross-sectional studies and prospective cohort studies were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well. Individuals were required to have suspected reactive arthritis at the time of testing, and therefore case–control studies, where people had a known disease status, were excluded. Studies were included even if they did not explicitly report diagnostic accuracy outcomes, as long as those outcomes could be calculated from the data presented.

A systematic search and a hand search of the reference lists of systematic reviews identified 6,769 references. The references were screened on their titles and abstracts and 36 studies were ordered for full text review. An additional 5 papers originally ordered for other diagnostic questions within this guideline were also identified as being potentially relevant and screened at the full-text level for this question.

Of these 41 studies, 36 were excluded as they did not meet the eligibility criteria for reasons such as an inappropriate study design (e.g. case reports, case–control) or the wrong study population (no suspected reactive arthritis at baseline). A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F. A total of 5 studies was included in the final review.

1 **4.6.1.1 Description of included studies**

2 The included studies are summarised in Table 19; full details are found in the evidence
3 tables (see Appendix E), with GRADE profiles reported in Appendix G.

4 **Table 19: Summary of included studies**

Study & location	Population	Reactive arthritis diagnostic criteria	Infection tests
Granfors 1980 Finland	37 people with acute infection with *Yersinia enterocolitica* serotype O:3, diagnosed by serological/bacteriological findings and clinical picture	Joint symptoms and subjective pain	Blood culture (ELISA for IgM, IgG and IgA antibodies to *Yersinia*)
Locht 1993 Sweden	29 people reporting joint symptoms after an outbreak of *Salmonella enteritidis* enterocolitis	Pain in a previously healthy joint at a well-defined anatomical location within the first 4 weeks after exposure	*S. enteritidis* stool culture
Mattila 1994 Finland	45 people reporting joint symptoms after an outbreak of *Salmonella bovismorbificans*	Development of synovitis (both swelling and tenderness) in a previously asymptomatic joint within the first weeks after the exposure in patients without another diagnosis or current inflammatory rheumatological diagnosis	Blood culture (*Salmonella*-specific IgM, IgA and IgG antibodies)
Toivanen 1987 Finland	104 people with acute infection with *Y. enterocolitica* serotype O:3, diagnosed by serological/bacteriological findings and clinical picture	Joint symptoms and subjective pain	Blood culture (ELISA for IgM, IgG and IgA antibodies to *Yersinia*)
Uotila 2011 Finland	45 people with swollen joints or sacroiliitis-like symptoms	Synovitis, tendinitis, enthesopathy, bursitis or probable sacroiliitis, with symptoms starting within 2 months of the outbreak	Faecal culture and antibodies against *Campylobacter*, *Salmonella* and *Yersinia*

5 **4.6.1.2 Minimal clinically important differences**

6 When considering individual factors in isolation, it was agreed in collaboration with the GDG
7 that a positive likelihood ratio of 2 would constitute significant diagnostic value. Therefore,
8 when interpreting the diagnostic accuracy results for single factors in isolation, something
9 would only be considered to have diagnostic value if the result was statistically significant at
10 the 95% confidence level, and the point estimate of the positive likelihood ratio was greater
11 than 2 (or equivalently, value at ruling out the disease if the negative likelihood ratio was less
12 than 0.5).

13 **4.6.2 Health economic evidence**

14 A literature search was conducted jointly for all review questions in this guideline by applying
15 standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
16 total of 9,970 references was retrieved, of which none were retained for this review question.
17 Health economic modelling was not prioritised for this review question.

4.6.3 Evidence statements

Low-quality evidence from 2 cross-sectional studies with 74 participants could not identify significant diagnostic value from analysing faecal or blood cultures for *Salmonella* following Salmonella outbreaks.

Low-quality evidence from 1 cross-sectional study with 45 participants could not identify significant diagnostic value from faecal or blood cultures for *Campylobacter*, *Salmonella* and *Yersinia*, following a *Campylobacter*, *Salmonella* and *Yersinia* outbreak.

Very low- to moderate-quality evidence from 2 prospective cohort studies with 138 participants found significant diagnostic value from negative IgA antibody to *Yersinia* in ruling out reactive arthritis at 6–16 months post-infection for people with suspected *Yersinia* reactive arthritis, but could not identify significant diagnostic value at earlier time points for positive IgA antibody to *Yersinia* results, or at any time points for IgM or IgG antibodies to *Yersinia* results.

4.6.4 Evidence to recommendations

Relative value of different outcomes	The guideline development group (GDG) agreed that the 2 key outcomes were whether testing for infection provided additional information to enable more accurate diagnosis of or ruling out of reactive arthritis, and whether the information would lead to changes in disease management. The GDG also agreed there were various subpopulations in which the diagnostic value of testing may be different, including people with a confirmed history of infection, people with a confirmed infection but known to have been exposed to an outbreak or incident case, and people with no known history of infection.
Trade-off between benefits and harms	The GDG noted that the evidence presented fell into two main categories. The first set of evidence related to people exposed to an infection outbreak, where there was not shown to be any diagnostic value to infection testing, with the GDG agreeing that the heightened antibody levels in everyone exposed to an outbreak meant there was likely to be little diagnostic value from testing. The second set of evidence was on the length of time the antibody response persists post-infection, and whether this differs between those who do or do not go on to develop reactive arthritis. Again, the GDG agreed there was little evidence to suggest value in testing, with only one of three antibodies tested showing a significant difference between people with and without reactive arthritis, and then only at certain time points. The GDG agreed that with the number of negative results found there could be little confidence this result was not the result of chance, and even if true did not provide sufficient evidence to suggest routine infection testing would be justified. The GDG agreed that, considering the generally negative pattern of evidence found for the diagnostic utility of infection testing in people with a history of infection, a recommendation not to routinely undertake such tests was justified. It was noted that this was in line with current practice, where it is rare to undertake testing specifically for the purpose of diagnosing reactive arthritis. They also agreed that, despite evidence being found for only three of the five infections specified in the protocol (none was identified for Shigella or Chlamydia), it was unlikely the pattern of results would be substantially different between differing gastrointestinal infections, and therefore a general recommendation was justified. The GDG also agreed that, since evidence was only identified in people either with a confirmed history of gastrointestinal infection, or who had been exposed to a large-scale outbreak, it would not be appropriate to

	make recommendations for the group of people without a history of gastrointestinal infection.
Economic considerations	No economic evidence was identified for this review question and original health economic modelling was not prioritised. The GDG agreed that, since it had made a negative recommendation, it was highly unlikely any substantial increase in resource use would result.
Quality of evidence	The GDG agreed there was generally a lack of evidence in this area, and the studies identified were mostly of a low quality. This was agreed to be mainly a result of the types of research studies commonly undertaken in this area. Many studies either recruit populations after an outbreak and only include those with a confirmed infection, or only recruit people known to have reactive arthritis and then undertaken infection testing. The lack of a comparator group in both of these study designs means it is not possible to extract diagnostic utility data from them, meaning that even though a considerable number of studies have been conducted, they have not provided much data useful to answering this question.
Other considerations	The GDG noted they had considered evidence solely for the diagnostic value of testing as part of the diagnosis of reactive arthritis, and not of the value of testing to enable treatment of the underlying infection itself. They agreed that in certain circumstances there may be good reasons why testing is beneficial to enable treatment of the underlying infection (particularly with chlamydia infections), and those decisions should not be impacted by the recommendations made here.

4.6.5   Recommendations

11. Do not routinely test for infective antibody status to diagnose reactive arthritis in people with a history of gastrointestinal infection

5 Pharmacological management

Patients with spondyloarthritis can have peripheral and/or axial inflammation. Treatment of peripheral and axial disease may need to be considered separately and may require different medicines for peripheral and axial disease. NICE technology appraisals guidance exists for the use of biological DMARDs for psoriatic arthritis, ankylosing spondylitis and non-radiographic axial spondyloarthritis. However, NICE guidance is not currently in place for people with peripheral spondyloarthritis without a diagnosis of psoriatic arthritis. There has been recent general acceptance of peripheral spondyloarthritis as a disease entity (with psoriatic arthritis as a diagnosis within this broader term).

Peripheral spondyloarthritis treatments can be local or systemic. Local treatments may include steroid injections to inflamed joints, tendon sheaths and enthesitis insertions. Topical non-steroidal anti-inflammatory drugs (NSAIDs) may have a role in superficial enthesitis or tendonitis. Systemic treatments include oral NSAIDs, corticosteroids, disease modifying anti-rheumatic drugs (DMARDs) and biological DMARDs. NSAIDs are frequently used in the symptomatic management of inflammatory joint symptoms, with the choice based on comorbidities and patient preference. Oral and intramuscular steroids may also occasionally be used to control flares of disease. Standard DMARDs, such as methotrexate, leflunomide and sulfasalazine, are used to manage peripheral spondyloarthritis, and have been shown to be effective in reducing joint damage. Where standard DMARDs are not effective in controlling disease activity, NICE guidance recommends the use of biological DMARD for psoriatic arthritis.

Axial spondyloarthritis treatments include oral NSAIDs, biological DMARDs and local corticosteroid injections (usually limited to local sacroiliac joint steroid injections). NSAIDs can be of particular use in axial spondyloarthritis, and have also been shown to reduce radiographic progression in axial inflammation in patients with raised inflammatory markers (Sieper 2005). Standard DMARDs are not routinely used in treating axial inflammation, but in a small proportion of patients may still be clinically indicated. Biological DMARDs have been shown to be of benefit in axial inflammation, and have become a mainstay of treatment in patients with more severe disease. No individual biological medicine has been shown to be more effective than any other in controlling axial inflammation. The choice of medicine should take into account cost, patient factors, such as frequency of injections and administration route, and clinical comorbidities, such as previous malignancy, infection risk or presence of demyelinating disease.

Many patients with spondyloarthritis have HLA-B27-related comorbidities, such as psoriasis, inflammatory bowel disease and uveitis. When choosing a standard or biological DMARD in spondyloarthritis, it is sensible to consider which therapies would provide optimal benefit to these comorbid conditions.

1 ## 5.1 Pharmacological interventions for axial symptoms of
2 ## spondyloarthritis

3 Review question 20

4 • What is the comparative effectiveness of the following pharmacological interventions for
5 management of axial spondyloarthritis:
6 ○ corticosteroids
7 ○ non-steroidal anti-inflammatory drugs
8 ○ standard disease-modifying anti-rheumatic drugs?

9 ### 5.1.1 Evidence review

10 The review focussed on identifying studies that fulfilled the conditions specified in Table 20. It
11 compared different non-steroidal anti-inflammatory drugs for the first-line treatment of axial
12 spondyloarthritis.

13 **Table 20: PICO table for use of NSAIDs in axial spondyloarthritis**

Population	People with spondyloarthritis
Intervention*	Aceclofenac
	Celecoxib
	Diclofenac
	Etoricoxib
	Fenoprofen
	Flurbiprofen
	Indomethacin
	Ketoprofen
	Meloxicam
	Naproxen (plain and enteric coated)
	Piroxicam
	Sulindac
	Tenoxicam
	Tolfenamic acid
Comparators	Another non-steroidal anti-inflammatory drug or placebo
Outcomes	Pain
	Discontinuation due to adverse effects
	Discontinuation due to lack of efficacy

14 **NSAIDs listed in the British National formulary (BNF), accessed May 2015, were of interest and data were*
15 *extracted if the dose used in the studies was within the recommended dose range specified in the BNF.*

16 For full details of the review protocol, see Appendix C. A systematic search identified 1,967
17 references. The references were screened on their titles and abstracts and 81 studies were
18 ordered for full text of which all were available.

19 Fifty six studies were excluded as they did not meet the eligibility criteria such as
20 inappropriate study design (e.g. non-randomised interventions) or ineligible clinical
21 population (e.g. reactive arthritis associated with Streptococcus infections). A detailed list of
22 excluded studies and reasons for their exclusion is provided in Appendix F. An additional 2
23 papers were identified through rerun searches at the end of the guideline

24 A total of 26 RCTs (Astorga 1987; Barkhuizen 2006; Batlle-Gualda 1996; Bird 1986; Burry
25 1980; Dougados 1999; Dougados 2001; Gibson 1980; Good 1977; Johnsen 1992; Juvakoski
26 1982; Khan 1987; Lomen 1986; Mayrhofer 1990; Nahir 1980; Pasero 1994; Rejholec 1980;
27 Schwarzer 1990; Shipley 1980; Sieper 2008; Sturrock 1974; Sydnes 1981; Tannenbaum
28 1984; van der Heijde 2005; Villa Alcazar 1996; Walker 2016) and 1 systematic review

(Guellec 2014) were included and provided data for the analyses of the three primary outcomes of interest.

5.1.1.1 Description of included studies

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G.

Pain

Twenty-four RCTs (Astorga 1987; Barkhuizen 2006; Batlle-Gualda 1996; Bird 1986; Burry 1980; Dougados 1999, Dougados 2001; Gibson 1980; Good 1977; Johnsen 1992; Juvakoski 1982, Khan 1987, Lomen 1986; Nahir 1980; Pasero 1994; Rejholec 1980; Schwarzer 1990; Shipley 1980; Sieper 2008; Sturrock 2008; Tannenbaum 1984; van der Heijde 2005; Villa Alcazar 1996; Walker 2016) comparing one NSAID with another NSAID or placebo were identified.

The RCTs included a total 3,214 participants diagnosed with ankylosing spondylitis (n = 3,164) and reactive arthritis (n=50) who received interventions for periods ranging from 1 week to 52 weeks. Where data were available, between 3% and 30% of participants were women and the mean age ranged from 34.0 years (SD=1.8) to 47.8 years (SD=8.8). Where stated, average disease duration ranged from 5.4 years (SD 5.4) to 16.5 years (No SD reported).

A systematic review (Guellec 2014) also contained data from 1 RCT comparing pain between continuous and on-demand NSAID therapy.

Discontinuation due to adverse effects

Nineteen RCTs (Barkhuizen 20066; Batlle Gualda 1996; Bird 1986; Burry 1980; Dougados 1999; Dougados 2001; Good 1977; Juvakoski 1982; Khan 1987; Lomen 1986; Mayrhofer 1990; Shipley 1980; Sieper 2008; Sturrock 1974; Sydnes 1984; Tannenbaum 1984; Van der Heijde 2005; Villa Alacazar 1996; Walker 2016) comparing 1 NSAID with another NSAID or placebo were identified. No data were identified for Tolfenamic acid.

The RCTs included a total 3,429 participants diagnosed with ankylosing spondylitis (n=3,379) and reactive arthritis (n=50) who received interventions for periods ranging from 1 week to 52 weeks. Where data were available, between 10% and 31% of participants were women and the mean age ranged from 34.0 years (SD=1.8) to 45.7 years (SD=11.7). Where stated, average disease duration ranged from 5.4 years (SD=5.4) to 16.5 years (No SD reported).

Discontinuation due to lack of efficacy

Fourteen RCTs (Barkhuizen 2006; Batlle Gualda 1996; Dougados 1999; Dougados 2001; Juvakoski 1982; Khan 1987; Lomen 1986; Mayrhofer 1990; Schwarzer 1990; Shipley 1980; Sieper 2008; Tannenbaum 1984; Van der Heijde 2005; Villa Alacazar 1996) comparing 1 NSAID with another NSAID or placebo were identified. No data were identified for Tolfenamic acid.

The RCTs included a total 2,932 participants diagnosed with ankylosing spondylitis (n=2,882) and reactive arthritis (n=50) who received interventions for periods ranging from 1 week to 52 weeks. Where data were available, between 10% and 31% of participants were women and the mean age ranged from 35.6 years (SD=1.3) to 45.4 years (SD=12.6). Where stated, average disease duration ranged from 6 years (No SD reported) to 13 years (SD=9).

5.1.1.2 Variations from protocol

Initially, the aim of this review was to assess the effectiveness of different pharmacological interventions in people with axial spondyloarthritis. However on GDG advice this was restricted to the effectiveness of individual NSAIDs in the management of the axial spondyloarthritis as NSAIDs were the only class of pharmacological intervention used as first line treatment in axial spondyloarthritis.

5.1.1.3 Minimal clinically important differences

A search in relation to axial spondyloarthritis of the Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted minimum clinically important difference thresholds for the primary outcomes in this review. Consideration was given to adopting a threshold percentage pain reduction that would be considered a clinically meaningful result. However, this was complicated by the fact that, since data from different pain scales were converted to a 0-100 scale for analysis, it is not clear that the same percentage reduction is equivalent across all these scales, which could lead to inconsistencies in the analysis (particularly when the network meta-analysis was conducted using an outcome of absolute reductions on the transformed scale, rather than proportional reductions). Further, it was agreed that any consistently measurable reduction in pain (i.e. one that could be shown to be a significantly greater reduction than random fluctuation) would be likely to be meaningful to patients, and therefore statistically significant differences in pain outcomes were agreed to be clinically meaningful. For other outcomes, where applicable, the GRADE default MID interval for dichotomous outcomes of (0.8 to 1.25) was used.

5.1.1.4 Network meta-analysis methods

The network meta-analysis for pain was undertaken by the NICE Technical Support Unit. Briefly, the pain scale outcomes from all studies were converted to a 0-100 scale (assuming an approximately symmetrical, unimodal distribution), and missing variances imputed from the distribution of reported standard deviations. A random effects consistency model was selected over a fixed effects consistency or random effects inconsistency model, based on the Deviance Information Criterion (DIC). Three Markov chain Monte Carlo (MCMC) chains with different initial values were run in WinBUGS 1.4.3, with convergence checked after 35,000 iterations. These values were then discarded as burn-in, and results based on a further 70,000 iterations run from that initial point. Full details of the methods are provided in Appendix J. An additional RCT, identified through update searches run at the end of the guideline, has been included in the analysis since this methods paper was written. Therefore, whilst the methodology is unchanged, final results reported in this guideline may differ slightly from those in the appendix.

Network meta-analyses for discontinuations were based on standard binomial logit models, with 35,000 iterations discarded as burn-in, and a further 70,000 iterations used to generate results. The DIC was used to select between fixed and random effects models, with results being downgraded for inconsistency if a random effects model was preferred. Indomethacin was chosen as the reference treatment for the analysis, as it was the most well-connected node in the network (i.e. there were more comparisons containing indomethacin than any other alternative).

5.1.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 6,450 references was retrieved, of which 2 were retained for this review question (Jansen et al., 2007; Jansen et al., 2010).

1   Jansen et al. (2007) compared etoricoxib with non-selective NSAIDs (nsNSAIDs) for first line
2   treatment for people with ankylosing spondylitis in the United Kingdom (UK). Treatment
3   effects only considered treatment withdrawals and no outcomes related to pain relief or
4   disease progression. They were based on etoricoxib trials that did not reflect the UK
5   ankylosing spondylitis population (78% female) and assumptions were required (including
6   data from wider arthritis populations) for the nsNSAIDs.

7   Drug costs were weighted by the average usage within each class, but neither the nsNSAIDs
8   used nor the average usage were detailed. Baseline utility data were mapped twice from
9   BASFI to SF-36 to utility, whilst utility changes (i.e. treatment effect) from changing
10 treatments due to lack of efficacy and adverse events (key outcomes) were partially
11 assumed.

12 Over 1 year, etoricoxib was found to be cost-effective (ICER £6,200/QALY) compared with
13 nsNSAIDs in 77% of iterations (at a £20,000/QALY threshold). However this was sensitive to
14 the nsNSAIDs drug costs – if the cheapest drug in each class were prescribed the cost
15 benefits of etoricoxib were reduced (ICER £16,200/QALY, cost-effective in 59% of iterations
16 at a £20,000/QALY threshold).

17 Using an expanded model, Jansen et al. (2010) compared etoricoxib with celecoxib,
18 naproxen and diclofenac for first line treatment for people with ankylosing spondylitis in the
19 UK. Treatment effects modelled were disease progression (but not pain) and were based on
20 a systematic review and a network meta-analysis (for which the included RCTs matched the
21 subset covering these comparators that were used in this review question). Adverse events
22 were taken from a wider arthritis population.

23 Drug costs were based on the most commonly prescribed drug within each class. Adverse
24 event costs were only applied in the first year. Utility data were mapped directly from BASFI
25 and BASDI, utility changes from changing treatments due to lack of efficacy and adverse
26 events were partially assumed.

27 Over a lifetime horizon, etoricoxib was found to dominate (produce fewer costs and more
28 QALYs) celecoxib, naproxen and diclofenac and was cost effective in 99.8% of iterations at a
29 threshold of £20,000/QALY.

30 Both papers were funded by the makers of etoricoxib. Neither paper modelled more than 1
31 switch of NSAID treatment; in both papers parameters for treatment intensification to anti-
32 TNFs relied on a number of assumptions (including cost, treatment effects and adverse
33 event rates).

34 Original health economic modelling was not prioritised for this review question.

5.1.3   Evidence statements

36 Network meta-analyses of 23 RCT's (pain), 18 RCT's (discontinuation due to adverse
37 events) and 14 RCT's (discontinuation due to lack of efficacy) found no statistically significant
38 difference between the NSAIDs examined, but that NSAIDs consistently outperform placebo
39 for pain control. Indomethacin appeared to have fewer discontinuations due to adverse
40 effects compared with some comparators (celecoxib (200mg), ketoprofen, piroxicam,
41 tenoxicam and placebo) at up to 1 year; etoricoxib, flurbiprofen and naproxen also appeared
42 better than placebo. Etoricoxib appeared to have statistically significantly fewer
43 discontinuations due to lack of efficacy compared with celecoxib (200mg), ketoprofen and
44 naproxen at up to 1 year.

45 Moderate- to high-quality evidence from 1 RCT of 215 people found slower radiographic
46 progression with continuous NSAID therapy compared to on-demand NSAID therapy, but
47 could not differentiate levels of pain. The same study found a higher incidence of depression

1   with continuous NSAID therapy compared to on-demand NSAID therapy, but could not
2   differentiate rates of serious adverse events.

3   Two partially applicable cost-utility analyses with potentially serious limitations found
4   etoricoxib to be cost-effective compared with the other NSAIDs modelled.

5   **5.1.4   Evidence to recommendations**

Relative value of different outcomes	The group considered that of the 3 prioritised outcomes, pain was the most important from the perspective of people with spondyloarthritis. This is supported by the idea that reduction in pain will also allow people with spondyloarthritis to engage more in physical activity with an attendant reduction in social isolation and also maintaining mobility and function over the longer term
	Discontinuation due to short-term (less than 52 weeks) adverse effects was considered to be of lower priority as these would generally be transient and would be likely to resolve once the NSAID had been stopped. However discontinuation due to lack of efficacy was considered important as the GDG noted that people with spondyloarthritis will only continue with treatment if the initial beneficial effect is maintained.
Trade-off between benefits and harms	Benefit was seen in terms of pain reduction with the majority of NSAIDs and the pain relief offered by NSAIDs outweighed the risk of adverse effects in shared decision making. The GDG was aware of a wider evidence base around long-term use of NSAIDs and an MHRA statement regarding increased risk of cardiovascular events with certain NSAIDs, but also of emerging evidence from UK registry data of 'reduced mortality risk' in this population. The GDG considered that this may be due to the fact that people who have less pain may be more able to engage in a range of physical activities which in turn may have positive impacts on long-term health.
	Compared with naproxen and celecoxib (200mg) the evidence presented for etoricoxib suggested some benefit in terms of reducing discontinuation due to lack of efficacy.
	The GDG agreed that the evidence presented reflected their own clinical experiences, in that most NSAIDs were similar, with perhaps etoricoxib and diclofenac performing slightly better than the others.
	The lack of longer term cardiovascular adverse event data was felt to be a limitation to decision making. It was noted that this was addressed by a different review question - here both the drug exposure and follow up were too short to enable the group to fully assess the risk of long term cardiovascular adverse events. It was noted that in recent years many people with spondyloarthritis had been switched from diclofenac (perceived to have the higher cardiovascular risk) to naproxen (perceived to have the lowest cardiovascular risk).
Economic considerations	Whilst the published economic evidence did not cover all the comparators in the decision space, the GDG agreed that the 2010 paper covered the main choices under consideration and was more applicable to their decision making. However, the paper was based on different outcomes (BASFI and BASDAI rather than pain VAS) to those presented and adverse event data was not specific to an ankylosing spondyloarthritis population.
	The GDG noted that etoricoxib appeared substantially more effective in the economic evidence than in the clinical evidence presented to them and this was driving the apparent cost-effectiveness of etoricoxib compared with other NSAIDs. Additionally, only applying adverse event costs in the first year of treatment was perceived as unrealistic given the long term exposure to NSAIDs and the uncertainty surrounding longer term cardiovascular adverse events.

	Both papers modelled the prescribing of proton-pump inhibitors following an adverse gastrointestinal event. Since the papers were published, the group noted that prescribing practice had changed and people would be prescribed PPIs with their initial NSAID.
	The GDG noted the extent of the limitations of the presented economic evidence and agreed that these limited their confidence in utilising the evidence when making recommendations.
Quality of evidence	The quality of the evidence was agreed to be moderate. The majority of studies were conducted before 2000 and the quality of reporting of either the methods or the outcomes was poor. However there was no clear evidence of bias in the included studies. The evidence was considered to be directly relevant as the population and the intervention in the included studies met the criteria stated in the review protocol. There was no significant inconsistency between the findings of the network meta-analysis when compared with the results from the pairwise analyses.
	The quality of the evidence was primarily downgraded as the rescaling of the pain outcome to a 0-100 scale for a number of studies relied on a number of assumptions regarding the original scales (symmetric, unimodal, same distributional shape). Additionally the GDG noted that the wide credible intervals and lack of significant differences between NSAIDs could be attributed to the imputed standard deviations in 9 out of 23 included papers
Other considerations	Age of studies – most studies were conducted prior to 2000 when indomethacin was the standard NSAID treatment option. This was before evidence of NSAIDs potentially increasing cardiovascular risk emerged. Since then, naproxen has become a more standard treatment. The group had limited experience of using tolfenamic acid and noted that only 1 small scale study was included in the evidence base for this NSAID.
	The NMA for drop outs due to adverse events did not specify or differentiate by different categories/types of adverse events.
	Ibuprofen was a notable gap in the included evidence base.
	Whilst not assessed here, the GDG agreed that there was still uncertainty as to the extent of the cardiovascular risk related to NSAID exposure and that NSAIDs may even have mortality protective properties in an axial spondyloarthritis population.
	The GDG also noted that the evidence presented did not report on individual short-term adverse effects, especially risk of cardiovascular events.
	The GDG expressed concern about the generalisability of the evidence identified to the wider population of people with spondyloarthritis, which would include those in whom NSAIDs may be contra-indicated. The fact that only 1 study included participants without a diagnosis of ankylosing spondylitis highlighted the paucity of evidence for axial symptoms in psoriatic arthritis and other non-AS conditions.
	Current prescribing practice was discussed and the GDG expressed concerns over 'blanket prescribing'. It also noted that some GPs may be unwilling to prescribe NSAIDs due to MHRA alerts and that people with spondyloarthritis may have difficulty obtaining prescriptions outside of specialist settings. The risk of litigation was raised as the GDG reported ongoing medical-legal cases over the prescribing of NSAIDs.
	The current clinical pathway for the pharmacological management of axial symptoms requires people with spondyloarthritis to have tried at least 2 NSAIDs before progressing to biological DMARDs. This raised the issue of sequencing of NSAIDs as the evidence presented did not report on this issue. However, the GDG noted that it was

common practice to assume that if 1 NSAID did not have an optimal response then another one may and that switching of NSAIDs was common clinical practice.

The GDG also noted that the experiences of people with spondyloarthritis were influenced by dosing regimens of individual NSAIDs. For example, whether they needed to take the dose once or twice a day. The presence or absence of comorbidities was also noted as influencing the choice of drug, including the potential prescribing of medicines other than NSAIDs.

5.1.5 Recommendations

12. Pharmacological management of axial spondyloarthritis

12.1. Offer NSAIDs to people with pain associated with axial spondyloarthritis.

5.2 Pharmacological management of peripheral spondyloarthritis

Review Question 21

- What is the comparative effectiveness of the following pharmacological interventions for the management of peripheral spondyloarthritis:
  - corticosteroids
  - non-steroidal anti-inflammatory drugs (NSAIDs)
  - standard disease-modifying anti-rheumatic drugs (DMARDs)?

5.2.1 Evidence review

The aim of this review was to compare the effectiveness of the specified first-line pharmacological interventions for peripheral spondyloarthritis.

Table 21: PICO table – management of peripheral spondyloarthritis

Population	People (aged 16 years and over) with a confirmed diagnosis of peripheral spondyloarthritis
Intervention	NSAIDs (Ibuprofen, Naproxen, Fenoprofen, flurbiprofen, Ketoprofen, Diclofenac, aceclofenac, Etodolac, Indomethacin, Meloxicam, Nabumetone, Phenylbutazone, Sulindac, Etoricoxib, Celecoxib)
	Corticosteroids (oral or injected) (Prednisolone, Prednisolone modified release, Betamethasone, Hydrocortisone [acetate], Solucorta (soluble), Methylprednisolone [acetate], Methylprednisolone sodium succinate (soluble), Triamcinolone acetonide, Triamcinolone hexacetonide)
	Standard DMARDs (Methotrexate, Sulfasalazine, intramuscular Gold, Leflunomide, Azathioprine, Ciclosporin)
Comparators	Any of the included interventions
Outcomes	Pain, adverse events, joint count, physical function, quality of life, imaging, composite measures, fatigue, CRP

For full details of the review protocol, please see Appendix C.

The search completed for this question yielded 1,992 references. These were screened on title and abstract and 131 full text papers ordered, 128 studies were excluded as they did not meet the eligibility criteria. A list of excluded studies with reasons for exclusion is provided in Appendix F. A total of 3 studies were included for this review question.

In accordance with the review protocol, only RCTs were included that had participants who were 16 years or older and had a confirmed diagnosis of spondyloarthritis. Where it was not stated it was assumed that those with psoriatic or reactive arthritis had peripheral disease.

5.2.1.1 Description of included studies

The included studies are reported in Appendix E and are described below, with GRADE profiles reported in Appendix G.

DMARD vs DMARD included studies

Two RCTs considered differing standard DMARD treatments (one compared ciclosporin with sulfasalazine and the other ciclosporin with methotrexate) in people with psoriatic arthritis and reported outcomes relevant to this review.

1 **Pain**

2 Pain outcomes in those with psoriatic arthritis, were included in 2 RCT studies comparing
3 differing standard DMARD treatments.

4 One 24 week study considered ciclosporin (initially 3mg/kg/day increased to a maximum of
5 5mg/kg/day with insufficient response) compared with sulfasalazine (1000mg/day for 1 week,
6 increased up to 2000mg/day, could be increased to 3000mg with insufficient response) and
7 both standard DMARDs compared with symptomatic therapy (NSAID/corticosteroids/
8 analgesics) in 99 participants. This study also reported on tender joint count and joint
9 pain/tenderness scores. Participants in this study had psoriatic arthritis with ≥3 swollen and
10 tender joints, 37% were male, with a mean (SD) arthritis duration of 1.9 (4.0) years in the
11 ciclosporin group, 2.7 (4.3) years in the sulfasalazine group and 2.0 (3.1) years in the
12 symptomatic therapy group. Participants in the standard DMARD groups in this study were
13 taking a stable NSAID dosage (Salavarani, 2001).

14 One 12 month study compared ciclosporin (3mg/kg/day, increased monthly to a maximum of
15 5mg/kg/day) with methotrexate (2.5mg every 12 hours, 3 doses, week 1, increased to a
16 maximum dose of 15mg/weekly) in 35 participants. This study reported on the number of
17 painful joints. Participants had psoriatic arthritis with ≥5 painful and/or swollen joints, 63%
18 were male, with a mean duration of arthritis of 9 years (range 1 to 32) in the ciclosporin group
19 and 8 years (range 1 to 21) in the methotrexate group. Participants were considered to not
20 be adequately controlled on NSAID and were on a stable NSAID dosage (Spardaro, 1995).

21 **Patient global assessment, physician global assessment**

22 Physician assessment of disease (via 100mm scale) was reported in 1 RCT of ciclosporin
23 compared with methotrexate, in 35 participants, who were considered to not be adequately
24 controlled on NSAID and were on a stable NSAID dosage (Spardaro, 1995).

25 **C reactive protein (CRP)**

26 CRP was reported in 2 RCTs. 1 study was of ciclosporin compared with sulfasalazine and
27 both standard DMARDs compared with symptomatic therapy (NSAID/corticosteroids/
28 analgesics) in 99 participants. People in the standard DMARD groups were taking a stable
29 NSAID dosage (Salavarani, 2001). The second study was of ciclosporin compared with
30 methotrexate, in 35 participants who were considered to not be adequately controlled on
31 NSAIDs and were on a stable NSAID dosage. At 12 months there was no significant
32 difference between the two treatment groups in CRP (Spardaro, 1995).

33 **Other measures**

34 The percentage of participants achieving the psoriatic arthritis response criteria from the
35 American College of Rheumatology (ACR), (ACR20, ACR50 and ACR70) were reported in 1
36 RCT. One study compared ciclosporin with sulfasalazine and both standard DMARDs
37 compared with symptomatic therapy (NSAID/corticosteroids/analgesics) (Salavarani, 2001).

38 **Adverse events**

39 Adverse events were reported in 1 RCT, which was of ciclosporin compared with
40 sulfasalazine and both standard DMARDs compared with symptomatic therapy
41 (NSAID/corticosteroids/analgesics). It contained 99 participants, and those in the standard
42 DMARD groups were taking a stable NSAID dosage. There were higher rates of
43 gastrointestinal adverse events with sulfasalazine (N=6, 19% compared with N=4, 11% with
44 ciclosporin) and higher rates of mild, reversible impaired renal function with ciclosporin
45 (N=10, 28% compared with N=1, 3% with sulfasalazine) and higher rates of neurological

disturbance (N=7, 19% compared with N=3, 9% with sulfasalazine) compared with the other groups (Salavarani, 2001).

NSAID vs NSAID included studies

One RCT considered differing NSAID treatments (ketoprofen and indomethacin) in people with reactive arthritis and reported outcomes relevant for this review (Juvakoski and Lassus, 1982).

Pain

One RCT, crossover study with 8 week study periods (1 week wash out), compared ketoprofen (200mg) compared with indomethacin (100mg), in 50 participants. This study reported on pain scores (details of scoring tool used not reported). Participants had reactive arthritis, 92% were male, mean duration of arthritis was 6 years (range 1 to 19).

Adverse events

This RCT reported small numbers of adverse events with ketoprofen (N=1 for each of diarrhoea, gastric pain, gastritis) and with indomethacin (N=1 for gastric pain, stomach pain, vertigo, dizziness and N=2 for headache).

Other possible outcomes

There were no studies identified that reported on quality of life, imaging or fatigue outcomes with the pharmacological interventions for peripheral spondyloarthritis.

5.2.1.2 Variations from protocol

The GDG agreed that the primary focus of the question should be on within-class comparisons between standard DMARDs.

The review protocol noted the possibility of including studies with a placebo comparator where these could be incorporated into a network meta-analysis. The review of the direct evidence did not suggest differences between the standard DMARD treatments in the included outcomes. There were a number of placebo comparator studies available that could be formed into network meta-analyses with limited numbers of standard DMARD studies in the linking arms (1 to 5 studies). It was considered that undertaking this analysis was unlikely to provide additional evidence that would aid decision-making and the development of recommendations by the GDG, therefore it was not undertaken.

5.2.1.3 Minimal clinically important differences

A search in relation to psoriatic arthritis and reactive arthritis of the Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted minimum clinically important difference thresholds for the outcomes in this review. Rheumatoid arthritis measures have been used previously within psoriatic arthritis assessment in clinical trials, though this has generally not been validated they have been used to distinguish placebo from treatment response (Mease, 2005). For pain outcomes, in those with chronic conditions (such as rheumatoid arthritis), there has been some consideration given to identifying the magnitude of pain reduction that could be considered clinically significant. Consensus from the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) has considered changes in pain scores in trials of patients with chronic pain (Dworkin, 2008). It was felt important to be consistent with the decisions made around MCIDs for the other treatment questions which used pain as a primary outcome measure. Further, it was agreed that any consistently measurable reduction in pain (i.e. one that could be shown to be a significantly greater reduction than random fluctuation) would be likely to be meaningful to

1 patients, and therefore statistically significant differences in pain outcomes were agreed to be
2 clinically meaningful. For other outcomes, where applicable, the GRADE default MID interval
3 for dichotomous outcomes of (0.8 to 1.25) was used.

4 ### 5.2.2 Health economic evidence

5 A literature search was conducted jointly for all review questions in this guideline by applying
6 standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
7 total of 9,970 references was retrieved, of which none were retained for this review question.
8 Health economic modelling was not prioritised for this review question.

9 ### 5.2.3 Evidence statements

10 **Pain**

11 Very low quality evidence found significantly larger decreases in overall pain with ciclosporin
12 compared with sulfasalazine in those with psoriatic arthritis. No difference was found
13 between the groups for tender joint count, swollen joint count and joint pain/tenderness
14 scores.

15 Very low quality evidence found no significant difference between the treatment groups for
16 ciclosporin and methotrexate in those with psoriatic arthritis for painful joints and swollen
17 joints.

18 Low quality evidence found no significant difference between the treatment groups with
19 ketoprofen compared with indomethacin, in those with reactive arthritis.

20 **Patient global assessment of disease, physician global assessment of disease**

21 Very low quality evidence found no significant difference between the treatment groups with
22 ciclosporin compared with sulfasalazine in physician global assessment of disease, in those
23 with psoriatic arthritis.

24 Very low quality evidence found no significant difference between the treatment groups for
25 ciclosporin and methotrexate in those with psoriatic arthritis for physician global assessment
26 of disease and for patient global assessment of disease.

27 **C reactive protein**

28 Very low quality evidence found no significant difference between the treatment groups with
29 ciclosporin compared with sulfasalazine or in ciclosporin compared with methotrexate in
30 CRP, in those with psoriatic arthritis.

31 Very low quality evidence found no significant difference between the treatment groups for
32 ciclosporin and methotrexate in those with psoriatic arthritis for CRP.

33 **Adverse events**

34 Very low quality evidence found small numbers of adverse events with ketoprofen (diarrhoea,
35 gastric pain, gastritis) and with indomethacin (gastric pain, vertigo, dizziness, headache) in
36 those with reactive arthritis.

37 ### 5.2.4 Evidence to recommendations

Relative value of different outcomes	This review presented evidence on the comparative effectiveness of DMARDs in the treatment and management of peripheral spondyloarthritis. The GDG agreed the importance of the included outcomes and

	requested the inclusion of swollen joints in the joint count outcomes as they considered this to be clinically relevant. The GDG considered that the dose of methotrexate given in 2 studies presented that could have been included was lower than would be prescribed for peripheral spondyloarthritis treatment currently in the UK. They therefore agreed that these studies did not provide useful evidence as the use of this low dose meant they could not be related to current clinical practice.
	The GDG agreed that the inclusion of placebo based studies in the network meta-analysis for DMARD therapy was unlikely to aid them in developing recommendations in this area. The limited available direct comparison evidence between individual DMARDs supported their clinical consensus that the efficacy of these drugs is similar in peripheral spondyloarthritis.
Trade-off between benefits and harms	The GDG considered that the included studies represented a substantially limited, very low quality evidence base. Nonetheless they agreed that this evidence supported their expert consensus and reflected current clinical practice that considers that there are not significant differences between the clinical effectiveness of different DMARD drugs used in the treatment of peripheral spondyloarthritis.
	The GDG discussed issues around the choice of DMARD and factors that may influence prescribing decisions. These included adverse effects, comorbidities and the use of methotrexate, ciclosporin and leflunomide in the management of psoriasis as well as psoriatic arthritis. The GDG discussed and noted the importance of considering information needs and personal preferences of those with spondyloarthritis in the decision making surrounding the prescribing of DMARDs.
	It was noted that the perception of people with peripheral spondyloarthritis on which treatments they preferred may be influenced by the information they receive on harms, co-morbidities, other disease manifestations (such as severity of psoriasis) or contraindications for particular DMARDs (such as contraindications associated with pregnancy or high levels of alcohol consumption).
Economic considerations	There was no economic evidence identified for this review question. The committee noted that drugs from this class are already commonly used in clinical practice, and therefore there are unlikely to be significant resource implications from the recommendations made.
Quality of evidence	Two studies were identified which considered different DMARD treatments compared with DMARDs in people with psoriatic arthritis. All relevant outcomes in these studies were rated as very low quality evidence using GRADE. It was noted that in both studies participants were taking NSAIDs and that these had been at a stable dose prior to study commencement. One further study was included that considered an NSAID treatment compared with another NSAID in people with reactive arthritis.
	The GDG agreed that these studies were eligible for inclusion, noting that they provide a very limited, poor quality evidence base, and that some of the reported interventions were not representative of UK practice (methotrexate dosage).
	The importance of establishing the number of joints affected by the peripheral spondyloarthritis to treatment decision making was discussed by the GDG. The GDG agreed that it is appropriate to consider the use of local or intramuscular corticosteroid injections for non-progressive monoarthritis rather than DMARD initiation in some cases. The GDG agreed that for peripheral polyarthritis, oligoarthritis and persistent or progressive monoarthritis DMARD therapy should be offered. In recognition of the limited evidence of differences in the efficacy of the individual DMARDs the GDG agreed the main considerations that may influence prescribing decisions, such as

	disease severity, associated comorbidities/complications, and lifestyle factors. The GDG noted that while first line DMARD therapy should be offered for peripheral spondyloarthritis, they thought NSAIDs may have a role in additional pain control and that they should be considered as adjunctive therapy. The GDG did not feel it was appropriate to use NSAIDs alone as a first line therapy, as they felt that DMARDs were necessary to target the disease activity in peripheral spondyloarthritis. They also considered that NSAIDs could be a useful adjunctive therapy in people receiving biological DMARDs, as they can help to manage symptoms (e.g. pain). Furthermore the GDG used their expertise and clinical experience to agree that intramuscular injections or short term oral steroid therapy may have a role as a short-term adjunctive treatment to standard or biological DMARDs in peripheral spondyloarthritis.
Other considerations	Due to the limited and low quality evidence base the GDG made consensus recommendations based on evidence where it was available and considered adequate, and their expert knowledge and clinical experience where evidence was lacking. They felt it was appropriate to make strong recommendations in some cases where there was limited or no evidence, if doing so supported best current practice. In particular, they agreed that DMARDs were established as being the first-line treatment of choice in treating peripheral spondyloarthritis, and this was well supported by evidence (e.g. Jones 2000, Pereda 2012) and clinical opinion. Whilst this evidence is not included in this review (as placebo comparisons were not included), it nonetheless was felt by the GDG to support the strength of the recommendations around DMARDs.

5.2.5 Recommendations

13. Pharmacological management of peripheral spondyloarthritis

13.1. Consider local corticosteroid injections as monotherapy for non-progressive monoarthritis.

13.2. Offer standard disease-modifying anti-rheumatic drugs (DMARDs) to people with:

- peripheral polyarthritis
- oligoarthritis
- persistent or progressive monoarthritis associated with peripheral spondyloarthritis.

13.3. When deciding which DMARD to offer, take into account:

- the person's needs, preferences and circumstances (such as pregnancy planning and alcohol consumption)
- comorbidities
- disease characteristics.

13.4. Consider NSAIDs as an adjunct to standard DMARDs or biological DMARDs to manage symptoms.

13.5. If NSAIDs do not provide adequate relief from symptoms, consider steroid injections (local or intramuscular) or short-term oral steroid therapy as an adjunct to standard DMARDs or biological DMARDs to manage symptoms.

5.2.6   Research recommendations

4.   What is the comparative effectiveness and cost-effectiveness of corticosteroids, NSAIDs and standard DMARDs for the management of peripheral spondyloarthritis, and is this effectiveness affected by differences in dose escalation protocols, frequency of monitoring or route of drug administration?

Why this is important

The committee noted that, whilst there are a number of randomised controlled trials comparing standard DMARDs with placebo for the management of peripheral spondyloarthritis, there is a lack of evidence comparing individual standard DMARDs to either NSAIDs or other standard DMARDs. This lack of evidence makes it difficult to optimise initial therapy, either by specifying specific drugs within the class or optimising dose, administration and monitoring protocols. There is therefore the need for randomised controlled trials looking at alternative drug, dosing and administration route alternatives for the pharmacological management of peripheral spondyloarthritis. These trials should include as outcomes measures both health-related quality of life (measured using the EQ-5D) and health service resource use, to enable the results to be used to assess the cost-effectiveness of the interventions.

5.3 Switching or augmenting pharmacological interventions for spondyloarthritis

Review Question 23

- When a first-line treatment has failed, what is the effectiveness of the following for managing spondyloarthritis:
    - switching to a different pharmacological intervention?
    - augmenting with a second pharmacological intervention?

5.3.1 Evidence review

The aim of this review was to identify where switching or augmenting pharmacological treatments may be effective where a first-line option has failed in people with either axial or peripheral spondyloarthritis.

Table 22: PICO table – switching or augmenting of pharmacological interventions

Population	People (aged 16 years and over) with a confirmed diagnosis of spondyloarthritis who did not respond to first-line therapy.
Interventions	- NSAIDs - Corticosteroids - Standard disease-modifying anti-rheumatic drugs - Biological DMARDs (in axial spondyloarthritis only)
Comparators	Each of the above when 1 first-line treatment option has failed, or as augmented therapy in combination with a first-line treatment.
Outcomes	Pain, adverse events, joint count/spinal mobility, physical function, quality of life, imaging, composite markers, inflammatory markers (ESR, CRP)

A single systematic search was conducted which identified a total of 1,103 references. The references were screened on their titles and abstracts and 132 studies were ordered for full text. All of these studies were excluded as they did not meet the eligibility criteria in the review protocol. A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F. Additionally the excluded studies lists in RQ20 and RQ21 were reviewed to consider whether there were any placebo comparator studies that met the review criteria for switching or augmenting treatment. This reference review identified 1 paper that has been included in this review (Fraser, 2005).

An additional paper was identified by the GDG that was published during guideline development (Coates, 2015).

5.3.1.1 Description of included studies

. Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G.

Methotrexate augmented with ciclosporin

One 12-month, multi-centre RCT reported on the addition of ciclosporin (to a maximum of 4 mg/kg/day) to methotrexate (compared with placebo and methotrexate) in participants with psoriatic arthritis who were considered to have had an incomplete response to a minimum of 15 mg of methotrexate weekly (lower if this dose was not tolerated) (Fraser 2005). Participants could be taking a stable dose of prednisolone, NSAID or both. There were 72 participants, with mean disease duration of 42.4 months (SD=41.9) for the placebo group

and 40.8 (SD=33.0) for the ciclosporin group. 76–79% of participants were also taking a stable dose of NSAIDs.

Tight control of inflammation in early psoriatic arthritis

One, 48 week, multi-centre, RCT reported on the use of a tight control treatment protocol for psoriatic arthritis compared with standard care in 8 secondary care centres in the UK (Coates et al., 2015). Two hundred and six participants with a diagnosis of psoriatic arthritis of less than 2 years' duration were included. Those previously treated for articular disease with DMARDs were excluded. The tight control protocol included review by the study physician every 4 weeks and escalation of treatment where psoriatic arthritis minimal disease activity criteria were not met.

5.3.1.2 Minimal clinically important differences

A search in relation to psoriatic arthritis and reactive arthritis of the Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted minimum clinically important difference thresholds for the outcomes in this review. It was felt important to be consistent with the decisions made around MCIDs for the other treatment questions which used pain as a primary outcome measure. Further, it was agreed that any consistently measurable reduction in pain (i.e. one that could be shown to be a significantly greater reduction than random fluctuation) would be likely to be meaningful to patients, and therefore statistically significant differences in pain outcomes were agreed to be clinically meaningful. For other outcomes, where applicable, the GRADE default MID interval for dichotomous outcomes of (0.8 to 1.25) was used.

5.3.2 Health economics evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which 1 was retained for this review question. Health economic modelling was not prioritised for this review question.

The included RCT reported by Coates et al. (2015) contained a brief description of a parallel economic analysis comparing intensive treatment with standard care in people with psoriatic arthritis. Details of the design, quality and results of the evaluation are tabulated in Appendix H.

5.3.3 Evidence statements

Ciclosporin and methotrexate

Very low-quality evidence found no differences between those treated with ciclosporin and methotrexate compared with those treated with methotrexate and placebo for the reductions in tender joint count, swollen joint count, CRP, patient global pain assessment, patient global disease activity and quality of life measures.

Tight control compared with standard care

Very low-quality evidence found significantly higher odds of achieving an ACR20, ACR50 or ACR70 response with tight control of inflammation in early psoriatic arthritis compared with standard care.

Very low-quality evidence found significantly higher rates of participants reaching HAQ MCID with tight control of inflammation in early psoriatic arthritis compared with standard care.

1 ### 5.3.3.1 Health economic evidence statement

2 A directly applicable cost–utility analysis with very serious limitations estimated that tight
3 control of inflammation in early psoriatic arthritis results in improved quality of life at
4 increased cost, with an ICER of £53,948 per QALY gained. The probability of a strategy of
5 tight control being cost effective in this population at a threshold of £20,000 per QALY is 7%.
6 With an assumption of a 25% reduction in total costs for both strategies and fewer
7 consultations in people who demonstrate adequate response to treatment on 2 consecutive
8 occasions of assessment, the ICER becomes £30,632.

9 ## 5.3.4 Evidence to recommendations

Relative value of different outcomes	The review question considered where first-line therapies for people with spondyloarthritis had failed and the subsequent switching or augmenting of drug treatments. The GDG agreed that this could also include the switching or augmenting of treatment following a tight control protocol as described in one arm of the TICOPA study (Coates et al., 2015).
	The GDG agreed the importance of the outcomes included in the review protocol and noted that response to treatment may be considered differently between people with spondyloarthritis and clinicians. It noted the importance of the individual's perception of efficacy being considered alongside measurable improvements in disease activity.
	The GDG discussed the outcome measures used in the TICOPA study. There was very low-quality evidence that the tight control group achieved the primary outcome (proportion of each group achieving ACR20) when compared with the standard care group. The outcomes were not considered directly relevant as they included rheumatoid arthritis research outcomes and axial spondyloarthritis measures. However, it was noted that these measures do include criteria that would be relevant to people with psoriatic arthritis.
Trade-off between benefits and harms	The GDG agreed that, if a person with spondyloarthritis also had other pre-existing conditions/symptoms which were successfully managed by the same classes of drugs, this should be taken into account when making decisions about whether augmentation or switching is appropriate. For example, people receiving treatment for psoriasis or inflammatory bowel disease may already be receiving drugs which can be prescribed for spondyloarthritis as well, so both conditions would need to be considered when choosing an optimal treatment plan. In addition, the GDG noted that some standard DMARDs may be less useful than others for the management of non-articular symptoms (e.g. sulfasalazine is less useful than methotrexate in the management of uveitis), which may influence the decision as to which drug to use, or whether a drug should be switched or augmented..
	The GDG noted that, in the case of NSAIDs, there would be no benefit in augmenting an NSAID that was not achieving a sufficient response with another NSAID. The GDG discussed switching between NSAIDs, noting that the network meta-analysis completed for review question 20 did not show substantial population-level differences in the comparative efficacy of different NSAIDs. The GDG agreed that individual responses will vary and therefore switching to a different NSAID may be worth considering where axial spondyloarthritis is not responding to the current choice. The GDG discussed the possibility that switching between NSAIDs could delay appropriate addition of treatment with biological DMARDs. After discussion it was agreed that there would be no substantial harm arising from switching NSAIDs in people who may benefit from biological DMARDs, as switching could be done concurrently with

	other clinical investigations such as pre-biologic screening (e.g. for TB) and the effectiveness of switching considered prior to prescribing or commencing biological DMARDs.
	The GDG discussed the possible harms of long term corticosteroid injections. No studies were identified in this review that directly considered long-term corticosteroid injections; however the GDG concluded that, though such studies had been carried out in non-inflammatory arthritis or other musculoskeletal conditions, this evidence should not be extrapolated to people with spondyloarthritis.
	The GDG discussed the possibility that people with extra-articular symptoms may be taking standard DMARDs for this condition and, whilst these may be effective for managing extra-articular considerations, they may not be effective for managing their spondyloarthritis. It was agreed that, where this occurs, the addition of a different standard DMARD should be considered for the spondyloarthritis symptoms.
	The GDG discussed the adverse event outcomes in the TICPOA trial. The differences between the groups were noted with adverse events reported more often by those in the tight control group. The GDG noted that the tight control group were reviewed more frequently throughout the study and that this, along with the differences in the treatments taken, may have had an impact on the reported adverse events.
Economic considerations	This review question was not prioritised for health economic modelling.
	No health economic evidence was found; however, the brief economic analysis from the TICOPA trial was considered on the advice of the GDG.
	The GDG discussed the economic evidence and noted that the quality of life improvements were small but the changes to costs were large. The incremental cost-effectiveness ratio (ICER) exceeded £50,000 per QALY, and the likelihood of a tight control strategy improving health at a cost lower than £20,000 per QALY was low. Sensitivity analysis around costs – both overall costs and the number of appointments necessary in participants reaching minimal disease activity – did not change this result.
	The economic analysis does not extend beyond the trial period, which the GDG agreed could have some important implications for the results, although without an exploration of the longer-term effects and costs of tight and standard control the expected changes cannot be established. Further economic modelling and/or longer-term follow-up of trial participants would enable comment to be made on the implications of tight control on quality of life and costs beyond the initial treatment escalation.
	The GDG acknowledged the forthcoming expiry of patents for some of the anti-TNFs and the emergence of 'biosimilars' which may facilitate a tight control strategy that approaches levels of cost effectiveness that are more acceptable. Without any detailed information on how patterns of prescribing and unit costs may change, however, the GDG felt unable to speculate on the impact this may have on the cost effectiveness of a tight control strategy like that assessed in TICOPA.
Quality of evidence	The GDG agreed that the available evidence was limited and of low quality. The GDG noted that in the included trial of methotrexate alone compared with methotrexate plus ciclosporin the methotrexate dose was not titrated up to the maximum tolerated dose before adding ciclosporin. This is unlikely to reflect standard UK practice where patients would have had their methotrexate dose optimised before augmenting treatment with a second standard DMARD. The GDG considered that regression to the mean may have occurred in

	the methotrexate-only group, as there was an improvement from baseline in both groups, and the participants had only been on methotrexate for 3 months at study initiation, so their treatment may not have truly 'failed'. It was noted that the trial included only a small number of participants and was sponsored by a pharmaceutical company.
For the TICOPA study, the GDG noted that there was a rapid escalation of treatment within the tight protocol arm and that this may have been more rapid than would be recommended clinically to allow sufficient time for some of the included treatments to reach effectiveness. The GDG agreed that, while comparison with standard care is an appropriate study design, the limited description of what constituted standard care within the study represents a risk of bias to the study outcomes. The GDG noted the importance of this study in considering new treatment protocols compared with standard care in an area where there is a lack of research; they agreed that some longer-term outcome data would be useful.
For peripheral spondyloarthritis the GDG noted the lack of evidence and discussed that, in practice, a standard DMARD that produces some effect but is considered not to have achieved an adequate response would be likely to be augmented with an additional standard DMARD. However the GDG agreed that, where there has been no response to a standard DMARD, then switching to an alternative standard DMARD would be appropriate. Therefore the recommendation includes consideration of either augmenting or switching between options in this class of drugs.
Other considerations	The GDG discussed the lack of evidence identified for this question and the low quality of the included studies. Therefore, when making their recommendations, the GDG considered this evidence and noted the pharmacological evidence presented in review questions 20 and 21. However, the recommendations made were primarily based on the GDG's expertise and experience.
The GDG discussed how a lack of response to a drug might be defined and agreed that changes to joints and (bio)markers could be indications of a response. It was noted that a person taking a single NSAID or DMARD may report no improvement despite a measurable response in some signs/symptoms, in which case augmentation may be more appropriate than switching. The GDG agreed that different drugs may need to be taken for different lengths of time before a clinical effect or improvement is detectable. SPCs should indicate the appropriate monitoring period and the optimal dose.
The GDG discussed the monitoring period used to determine whether one drug was effective before switching to/augmenting with a second drug and agreed that this would have to allow for the different lengths of time that different drugs within a class may take to show an effect. Therefore, the GDG added to the recommendations the time frames that a response could be expected to occur in as these are different between NSAID and standard DMARD treatments.
The GDG noted that all of the identified evidence came from studies of people with psoriatic arthritis. For the purposes of this question, the GDG felt it was appropriate to extrapolate the findings, combined with their own clinical experience, to make recommendations applicable to all forms of peripheral arthritis.

5.3.5 Recommendations

14. Second-line pharmacological management of axial spondyloarthritis

14.1. If an NSAID taken at the maximum tolerated dose for 2–4 weeks does not provide adequate pain relief, consider switching to another NSAID.

15. Second-line pharmacological management of peripheral spondyloarthritis

> **15.1. If a standard DMARD taken at the maximum tolerated dose for at least 3 months does not provide adequate relief from symptoms, consider switching to or adding another standard DMARD.**

> **15.2. If extra-articular disease is adequately controlled by an existing standard DMARD but spondyloarthritis is not, consider adding another standard DMARD.**

5.3.6 Research recommendations

> **5. When first-line treatment for spondyloarthritis has failed, what is the most effective and cost-effective ordering of systemic biological disease-modifying anti-rheumatic drugs to treat with and does this ordering change based on particular patient characteristics?**

Why this is important

Only a limited amount of low-quality evidence was found looking at the effectiveness of switching or augmenting treatment when first-line treatment is not providing adequate symptom control, and therefore it was only possible to make very general, class level, recommendations. Well conducted RCTs comparing different possible alternatives for second-line treatment, and looking at whether the optimum second-line treatment differs based on patient characteristics, would enable more specific and individually tailored treatment choices to be made in the future.

5.4 Biological DMARDs for spondyloarthritis

Review questions 24, 25 and 26

- What is the effectiveness of systemic biological disease-modifying anti-rheumatic drugs for managing symptoms of enteropathic arthritis?
- What is the effectiveness of systemic biological disease-modifying anti-rheumatic drugs for managing symptoms of reactive arthritis?
- What is the effectiveness of systemic biological disease-modifying anti-rheumatic drugs for managing symptoms of undifferentiated spondyloarthritis, excluding non-radiographic ankylosing spondylitis?

5.4.1 Evidence review

The aim of these reviews was to assess the effectiveness of biological DMARDs for the second-line management of symptoms (axial and peripheral) of reactive arthritis, enteropathic spondyloarthritis and undifferentiated spondyloarthritis. Evaluation of biological DMARDs for axial spondyloarthritis and psoriatic arthritis was outside of the scope of this guideline, as NICE guidance can be found in existing or forthcoming NICE Technology Appraisals, which are either cross-refered to or incorporated within this guideline.

The review focused on identifying studies that fulfilled the conditions specified in Table 23, Table 24 and Table 25.

Table 23: PICO table for question 24: biological DMARDs for enteropathic spondyloarthritis

Population	People (aged 16 years and over) with a confirmed diagnosis of enteropathic spondyloarthritis
Interventions	Biological DMARDs, to include: abatacept, adalimumab, anakinra, cacakinumab, certolizumab pegol, etanercept, golimumab, infliximab, natalizumab, rituximab, secukinumab, tocilizumab, ustekinumab Small molecule drugs: apremilast, tofacitinib
Comparators	Any of the above, plus placebo, or other classes of systemic drugs used to treat this group (NSAIDs, DMARDs, corticosteroids)
Outcomes	Pain, adverse events, joint/spinal mobility, physical function, quality of life, imaging, composite measures

Table 24: PICO table for question 25: biological DMARDs for reactive arthritis

Population	People (aged 16 years and over) with a confirmed diagnosis of reactive spondyloarthritis
Interventions	Biological DMARDs, to include: abatacept, adalimumab, anakinra, cacakinumab, certolizumab pegol, etanercept, golimumab, infliximab, natalizumab, rituximab, secukinumab, tocilizumab, ustekinumab Small molecule drugs: apremilast, tofacitinib
Comparators	Any of the above, plus placebo, or other classes of systemic drugs used to treat this group (NSAIDs, DMARDs, corticosteroids)
Outcomes	Pain, adverse events, joint/spinal mobility, physical function, quality of life, imaging, composite measures

1 **Table 25: PICO table for question 26: biological DMARDs for undifferentiated
2 spondyloarthritis**

Population	People (aged 16 years and over) with a confirmed diagnosis of undifferentiated spondyloarthritis (excluding non-radiographic axial spondyloarthritis)
Interventions	Biological DMARDs, to include: abatacept, adalimumab, anakinra, cacakinumab, certolizumab pegol, etanercept, golimumab, infliximab, natalizumab, rituximab, secukinumab, tocilizumab, ustekinumab Small molecule drugs: apremilast, tofacitinib
Comparators	Any of the above, plus placebo, or other classes of systemic drugs used to treat this group (NSAIDs, DMARDs, corticosteroids)
Outcomes	Pain, adverse events, joint/spinal mobility, physical function, quality of life, imaging, composite measures

3   For full details of the review protocols please see Appendix C. Evidence tables for included
4   studies can be found in Appendix E, with GRADE profiles reported in Appendix G.

5   Randomised controlled trials were considered to be the highest quality evidence available to
6   answer this question and are graded as high in a GRADE framework if conducted and
7   reported well.

8   A systematic search and a hand search of the reference lists of systematic reviews identified
9   3,478 references. The references were screened on their titles and abstracts and 59 studies
10 were ordered for full text, of which 55 were available. A further 4 studies were ordered during
11 an update of the review, of which 2 were available.

12 Fifty-six studies were excluded because they did not meet the eligibility criteria such as
13 inappropriate study design (for example, studies lacking a comparison group) or non-primary
14 studies (for example, systematic reviews and editorials). A detailed list of excluded studies
15 and reasons for their exclusion is provided in Appendix F. One RCT was included
16 (Paramarta et al., 2013).

17 ### 5.4.1.1   Variations from protocol

18 The 3 review protocols each specified consideration of a single indication for biological
19 DMARDs. No studies were identified which reported on outcomes separately by eligible
20 indication. However, 1 study limited its remit to a mixed population of people with peripheral
21 spondyloarthritis (explicitly excluding people with ankylosing spondylitis or psoriatic arthritis)
22 and this study was included. The majority of participants in this study had undifferentiated
23 spondyloarthritis.

24 ### 5.4.1.2   Description of included studies

25 One study was identified (Paramarta et al., 2013) which compared adalimumab with placebo
26 over a 12-week period in people with peripheral spondyloarthritis, excluding people with
27 ankylosing spondylitis or psoriatic arthritis. At randomisation there were 20 people per study
28 arm.

29 In the adalimumab group, 15 people had undifferentiated spondyloarthritis (75%), 4 had
30 reactive arthritis (20%) and 1 had enteropathic spondyloarthritis (5%). There were 9 men and
31 11 women in this group, with an average age of 41.5 years (SD 12.8) and mean disease
32 duration of 7.9 years (SD 9.3). Concomitant medication use was as follows: 13 (65%) were
33 taking NSAIDs, 5 (25%) were taking methotrexate, 7 (35%) were taking sulfasalazine and 1
34 (5%) had previous history of anti-TNF treatment. None were receiving corticosteroids.

1   In the placebo group, 17 people had undifferentiated spondyloarthritis (85%), none had
2   reactive arthritis and 3 had inflammatory bowel disease-related spondyloarthritis (15%).
3   There were 12 men and 8 women in this group, with an average age of 44.4 years (SD 11.1)
4   and mean disease duration of 6.7 years (SD 6.2). Concomitant medication use was as
5   follows: 14 (70%) were taking NSAIDs, 6 (30%) were taking methotrexate, 4 (20%) were
6   taking sulfasalazine, 2 (10%) had previous history of anti-TNF treatment and 2 (10%) were
7   receiving corticosteroids.

8   Study participants received subcutaneous injections of either 40 mg of adalimumab or
9   placebo every other week for 12 weeks. The primary study endpoint was improvement in
10   patient's global assessment of disease activity at week 12. There then followed an open-
11   label extension for weeks 12 to 24 in which participants in both study arms received
12   adalimumab.

13 **5.4.1.3   Minimal clinically important differences**

14   A search in relation to axial spondyloarthritis, psoriatic arthritis and reactive arthritis of the
15   Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted
16   minimum clinically important difference thresholds for the outcomes in this review. It was felt
17   important to be consistent with the decisions made around MCIDs for the other treatment
18   questions which used pain as a primary outcome measure. Further, it was agreed that any
19   consistently measurable reduction in pain (i.e. one that could be shown to be a significantly
20   greater reduction than random fluctuation) would be likely to be meaningful to patients, and
21   therefore statistically significant differences in pain outcomes were agreed to be clinically
22   meaningful. For other outcomes, where applicable, the GRADE default MID interval for
23   dichotomous outcomes of (0.8 to 1.25) was used.

24 **5.4.2   Health economics evidence**

25   A literature search was conducted jointly for all review questions in this guideline by applying
26   standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
27   total of 9,970 references was retrieved, of which none were retained for this review question.
28   Health economic modelling was not prioritised for this review question

29 **5.4.3   Evidence statements**

30   Very low- to low-quality evidence from a single RCT identified an improvement in swollen
31   joint count, BASDAI score and ESR levels in people with peripheral spondyloarthritis who
32   received adalimumab, compared with those who received placebo.

33   Very low-quality evidence from a single RCT identified no difference in tender joint count,
34   CRP levels, quality of life (measured by HAQ-DI and HUI-3) or the number of people
35   experiencing any or severe adverse events in people with peripheral spondyloarthritis who
36   received adalimumab, compared with those who received placebo.

37 **5.4.4   Evidence to recommendations**

Relative value of different outcomes	The GDG noted the relevance of including this review on the biological DMARDs for indications that are not covered by current NICE Technology Appraisal guidance.
	The GDG agreed the importance of the pre-specified outcomes as these reflected the potential benefits of biological DMARDs and the adverse events that may be associated with their use.
Trade-off between benefits and harms	The GDG noted that there is extensive experience in using biological DMARDs for many indications, including other spondyloarthritides. It agreed that, while it would be difficult to guess how effective the drugs would be for treating reactive, enteropathic or undifferentiated spondyloarthritis on the basis of this experience, the adverse effects

	of the agents would be more likely to generalise across indications. Therefore, the GDG agreed it should be cautious about recommending treatments with known harms and uncertain benefits.
Economic considerations	No economic evidence was presented. The GDG were aware that biological DMARDs are comparatively expensive technologies, and would have required convincing evidence as to their effectiveness to recommend their use.
Quality of evidence	One paper was identified for inclusion in the evidence review. The GDG agreed with the GRADE quality review of the included evidence which considered that for all of the included outcomes the quality of evidence was very low. The GDG noted that this RCT study included 20 participants in each arm and compared adalimumab with placebo. The GDG agreed the evidence statements reflected the study findings.
	The GDG discussed whether the definition of undifferentiated spondyloarthritis used in the paper would overlap with the definition now used for non-radiographic axial spondyloarthritis; if so the relevance of this paper to the review question could be considerably reduced as the population covered in the paper would be a mixture of eligible and non-eligible people according to the NICE guideline review protocol. The GDG agreed that the paper should be included in the evidence review but that it provided very low-quality indirect evidence. The RCT provided very limited evidence in relation to the difference found from baseline in swollen joint counts in people with peripheral spondyloarthritis. The GDG further noted that around 75% of patients in this study also had axial symptoms.
	The GDG discussed whether it would be possible to make any recommendations in this area. It considered the very limited evidence identified and whether it would be appropriate to consider biological DMARDs for enteropathic, reactive and undifferentiated spondyloarthritis as an extrapolation from the NICE Technology Appraisals (TAs) which considered the use of biological DMARDs in those with ankylosing spondylitis and/or non-radiographic axial spondyloarthritis (TA383) and psoriatic arthritis (TA199, TA220 and TA340).
	The GDG discussed the issues arising from this possible extrapolation of existing evidence, including ensuring that it considered the possible benefits and harms of these potential treatments for these conditions. It noted that these are potentially expensive therapies. It also noted that, in the TA guidance, it is clear when treatment should be commenced and how the response can be assessed with scoring tools. It discussed and agreed that, for the included conditions in these questions, these decisions would be less clear, and with peripheral arthritis an adequate response may be difficult to assess. It noted that, for the existing TAs there was evidence available, while, for the conditions under consideration here, the evidence base is very poor.
	The GDG noted the need to review diagnoses that patients have received in the past, as classifications/diagnostic criteria have changed. The GDG noted that the number of people presenting with the conditions under consideration may be quite small and consideration of individual circumstances may be more appropriate, for example through an individual funding request process. The GDG concluded that there is poor evidence in this area and that the benefits of biological DMARDs, including when to commence and how to assess response, could not be judged from the available evidence or confidently extrapolated from the existing NICE TA guidance. Therefore the group agreed that it could not make recommendations for this review question.
Other considerations	Due to the lack of relevant evidence available in this area, the GDG

opted to make a research recommendation.

5.4.5 Recommendations

No recommendations were made.

5.4.6 Research recommendations

6. What is the effectiveness and cost-effectiveness of biological DMARDs in people with persistent peripheral spondyloarthritis (excluding psoriatic arthritis) or undifferentiated spondyloarthritis?

Why this is important

Although there have been trials conducted of biological DMARDs for psoriatic arthritis, which have led to positive recommendations in NICE technology appraisals, no such good quality evidence exists in enteropathic arthritis, reactive arthritis or undifferentiated spondyloarthritis. The substantial side effects possible with biological DMARDs, and their significant cost, means it is difficult to justify offering them to these groups without good evidence of efficacy. There is therefore the need for randomised controlled trials, with a sufficient sample size to identify possible benefits, in these 3 populations. If trials were to recruit participants from multiple spondyloarthritis subpopulations, results should be clearly stratified by diagnosis to enable any differences in benefits or harms between the groups to be identified. These trials should include as outcomes measures both health-related quality of life (measured using the EQ-5D) and health service resource use, to enable the results to be used to assess the cost-effectiveness of the interventions.

5.5 Long-term antibiotics for reactive arthritis

Review Question 19

- What is the effectiveness of long-term (4 weeks or longer) treatment with antibiotics for first-line management of reactive arthritis compared with standard treatment?

5.5.1 Evidence review

The aim of this review was to assess the effectiveness of long-term antibiotic treatment for first-line management of symptoms of reactive arthritis.

The review focused on identifying studies that fulfilled the conditions specified in Table 26.

Table 26: PICO table for question 19: antibiotics

Population	Patients with reactive arthritis, with or without confirmation of causative infectious agent. Patients with a confirmed diagnosis of reactive arthritis triggered by microbes other than *Campylobacter, Chlamydia, Salmonella, Shigella* or *Yersinia* were excluded.
Interventions	Antibiotics for a minimum of 4 weeks
Comparators	Standard treatment to include placebo, and non-antibiotic first line therapies being used to manage arthritis (e.g. standard DMARDs, NSAIDs, corticosteroids).
Outcomes	Pain, adverse events, joint count, physical function, sacroiliitis imaging, inflammatory markers (ESR, CRP), fatigue

For full details of the review protocol please see Appendix C.

Randomised controlled trials were considered to be the highest quality evidence available to answer this question and were graded as 'High' in a GRADE framework if conducted and reported well.

A systematic search and a hand search of the reference lists of systematic reviews identified 1,373 references. The references were screened on their titles and abstracts and 27 studies were ordered for full text of which all were available.

Sixteen studies were excluded as they did not meet the eligibility criteria such as inappropriate study design (e.g. non-randomised interventions) or ineligible clinical population (e.g. reactive arthritis associated with Streptococcus infections). A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F.

A total of 10 RCTs and 1 secondary study providing long-term follow up was included.

5.5.1.1 Variations from protocol

The guideline scope specifies that reactive arthritis cases which are HLA-B27 associated were within remit. In many cases, studies measured prevalence of HLA-B27 within their trial populations and found that some participants were not positive for the marker. These studies and participants were nonetheless retained providing that the triggering microbial agent had not been confirmed to be outside of the pre-specified inclusion list (*Campylobacter, Chlamydia, Salmonella, Shigella* or *Yersinia*).

One of the included papers was a long-term follow-up study of a trial which was included in the main analysis. The results of this long-term study were analysed separately.

Prior to analysis, clinically relevant subgroups were identified (urogenital infection cases and gastrointestinal infection cases) and subgroup analysis was performed where data were available.

5.5.1.2 Description of included studies

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G.

All interventions and eligible triggers of reactive arthritis

Ten RCTs[a] were identified which compared long term antibiotic treatment (minimum of 4 weeks) with placebo in patients with confirmed or suspected reactive arthritis triggered by an eligible microbial infection. They included a total of 524 participants who received interventions for periods ranging from 4 weeks (Whaley 1969) to 12 months (Wakefield 1999). Where data were available, 39.7% of participants were women and the mean age ranged from 25.5 to 44.2 in the intervention groups and 22.4 to 49.0 in the placebo groups. Studies ranged from infection-specific populations (Carter 2010; Putschky 2006; Hoogkamp-Korstanje 2000) to mixed populations with either confirmed or suspected triggers, or a mixture of the two. Where stated, average duration of reactive arthritis ranged from a median of 5 weeks to a mean of 10.4 years in the intervention groups, and median 4 weeks to mean of 14.2 years in the placebo groups.

Five studies used ciprofloxacin (Sieper 1999, Toivanen 1993, Wakefield 1999, Yli-Kerttula 2000, Hoogkamp-Korstanje 2000), one used azithromycin (Kvien 2004), one lincomycin (Whaley 1969) and one used doxycycline alone (Putschsky 2006). Two studies used combination therapies in comparison to placebo: one assigned ofloxacin plus roxithromycin (Kuuliala 2013) and the other used either doxycycline plus rifampicin or azithromycin plus rifampicin (Carter 2010).

Urogenital triggers only

Two studies (Carter 2010; Putschky 2006) presented outcomes on urogenital infection-triggered reactive arthritis (specifically confirmed Chlamydia infection). Of the total study population (n=74), 47.3% were women and the mean age ranged from 42.6 to 44.2 in the intervention groups and 40.5 to 49.0 in the placebo groups. Duration of disease ranged from 17.1 months (range 2–24) to 10.4 years (SD=12.1) in the intervention groups, and 16.0 months (range 5–49) to 14.2 years (SD=14.2) in the placebo groups. One study (Putschky 2006) assigned participants to either doxycycline or placebo for 4 months. The other (Carter 2010) allocated participants to either triple-placebo, or doxycycline plus rifampin plus placebo, or azithromycin plus rifampin plus placebo, for 3 months. This study grouped together the 2 active intervention arms during analysis.

Gastrointestinal triggers only

One study (Hoogkamp-Korstanje 2000) presented outcomes solely on gastrointestinal infection-triggered reactive arthritis. All participants in this study had diagnosed Yersinia infection and were allocated to receive either ciprofloxacin or placebo for 3 months. Of the 18 participants, the mean age in the ciprofloxacin group was 33 (range 18–52) and in the placebo group was 45 (range 26–72); 44.4% of participants were women. Mean disease duration was 1.9 years (SD=1.4) in the ciprofloxacin group and 2.0 years (SD=1.5) in the placebo group.

One study (Sieper 1999) involved a population with a mixture of microbial triggers who had received diagnoses of either reactive arthritis or undifferentiated oligoarthritis. There were 39 participants with reactive arthritis triggered by either Yersinia or Salmonella for whom separate outcome data were reported.

---

a Carter 2010, Hoogkamp-Korstanje 2000, Kuuliala 2013, Kvien 2004, Putschsky 2006, Sieper 1999, Toivanen 1993, Wakefield 1999, Whaley 1969, Yli-Kerttila 2000.

Long-term secondary follow up

One secondary study (Yli-Kerttula 2003) described long-term follow up of an included randomised controlled trial (Yli-Kerttula 2000). Of the 71 participants of the original trial, 69 were contacted and invited to participate in follow-up investigations either by telephone (n=16) or via a face-to-face appointment at the clinic (n=53). Analysis was limited to the clinic-attending participants who had an average age of 36.8 (SD=12.4). 43.4% were women, and the average duration of disease in the index episode was 34.9 days (SD=23.7). The original intervention of ciprofloxacin vs placebo had lasted for 3 months; this follow-up study provided 4–7 years post-intervention follow up. 84.9% of patients in the follow-up analysis were HLA-B27 positive; this dropped to 18.9% of those who were identified as having chronic disease, all of whom had received placebo during the index episode.

5.5.1.3 Minimal clinically important differences

A search in relation to reactive arthritis of the Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted minimum clinically important difference thresholds for the outcomes in this review. It was felt important to be consistent with the decisions made around MCIDs for the other treatment questions which used pain as a primary outcome measure. Further, it was agreed that any consistently measurable reduction in pain (i.e. one that could be shown to be a significantly greater reduction than random fluctuation) would be likely to be meaningful to patients, and therefore statistically significant differences in pain outcomes were agreed to be clinically meaningful. For other outcomes, where applicable, the GRADE default MID interval for dichotomous outcomes of (0.8 to 1.25) was used.

5.5.2 Health economics evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

5.5.3 Evidence statements

All interventions and eligible triggers of reactive arthritis

Low- to very low-quality evidence showed no effect of long-term antibiotic treatment on painful or tender joints or arthralgia, swollen joints, general pain, stiffness, ESR, CRP or fatigue in RCTs of any antibiotic compared with placebo in people with HLA-B27–associated reactive arthritis.

Urogenital triggers only

Low- to very low-quality evidence showed no effect of long-term antibiotic treatment on painful or tender joints or arthralgia, swollen joints, stiffness, general pain, ESR, CRP or fatigue in randomised controlled trials of any antibiotic compared with placebo in people with urogenital infection–associated reactive arthritis.

Gastrointestinal triggers only

Very low-quality evidence showed no effect of long-term antibiotic treatment on painful or tender joints or arthralgia in randomised controlled trials of any antibiotic compared with placebo in people with gastrointestinal infection–associated reactive arthritis.

1   **Long-term secondary follow up**

2   Very low-quality evidence showed a reduction in late clinical findings of spondyloarthritis
3   following long-term antibiotic treatment during the index episode in people receiving
4   ciprofloxacin, compared with people receiving placebo in HLA-B27–associated reactive
5   arthritis.

6   Very low-quality evidence showed no effect of long-term antibiotic treatment on radiographic
7   findings, MRI findings, or ESR in a long-term follow-up study of an RCT of ciprofloxacin
8   compared with placebo in people with HLA-B27–associated reactive arthritis.

9   **5.5.4   Evidence to recommendations**

Relative value of different outcomes	The most important patient outcomes were considered to be joint pain and swelling.
	The GDG also noted that CRP can be important if elevated as it can indicate the level of infection/inflammation and, if reduced (in people who have had high CRP levels), is a good indicator of response to treatment.
Trade-off between benefits and harms	The GDG discussed concerns regarding the over-prescription of long-term antibiotics and noted the current drive to reduce the overall use of antibiotics to prevent resistance issues and *Clostridium Difficile*.
	The possible benefits of the long-term use of antibiotics are to prevent long-term problems associated with reactive arthritis. However it was noted that some infections responsible for reactive arthritis do not routinely require antibiotics (e.g. Salmonella).
	The evidence review showed no overall significant benefits and harms associated with the long-term use of antibiotics in outcomes such as fatigue, pain, CRP or adverse events.
Economic considerations	No health economic evidence was found and this review question was not prioritised for health economic modelling. The GDG agreed that, since it had made a negative recommendation, it was highly unlikely any substantial increase in resource use would result.
Quality of evidence	The GDG agreed that the evidence base was generally of low to very low quality due to methodological issues within and across the included studies.
	The GDG also raised the issue of comparability between studies with different disease duration (such as 16 months vs. 14.2 years) with the expectation that newly diagnosed or more recently diagnosed patients should show greater benefit. The GDG also discussed the possible confounding effects of additional medications within the long-term follow-up longitudinal study.
Other considerations	The majority of patients with acute reactive arthritis settle over 6-12 months but up to 20% of patients may develop persistent disease. Therefore follow-up may be required for a year to ensure patients do not have chronic persistent disease which will require on-going management. There is no evidence to support the use of long term antibiotics to improve prognosis. Therefore acute infection should be treated appropriately (which may involve antibiotics for some infections e.g. chlamydia), but there does not seem to be any evidence of long-term benefits for reactive arthritis particularly in terms of severity of disease.

1 **5.5.5 Recommendations**

2 **16. After treating the initial infection, do not offer long-term (4 weeks or longer)**
3 **treatment with antibiotics solely to manage reactive arthritis caused by a**
4 **gastrointestinal or genitourinary infection.**

6 Non-pharmacological management

Spondyloarthritis can cause joint, tendon and spinal pain, joint swelling and stiffness, fatigue, physical disability, and significantly reduce a person's physical fitness, quality of life and psychological wellbeing. In addition, spondyloarthritis is associated with a number of other potential complications. Management includes pharmacological and non-pharmacological strategies to help prevent or manage the significant impact on the person's physical and psychological wellbeing.

There are a number of non-pharmacological interventions that are commonly used or considered to be of potential benefit in managing problems caused by spondyloarthritis. In developing this guideline interventions including manual therapy, exercise programmes, hydrotherapy, acupuncture and the provision of physical aids have been evaluated. Their effectiveness in preventing, improving or managing symptoms, physical impacts, complications and consequences of inflammatory diseases are important to establish in order to guide practice, service delivery and help people with spondyloarthritis self-management their disease.

Manual therapy encompasses a number of approaches used in managing pain and joint and tendon dysfunction. It involves therapist-performed or therapist-assisted techniques to facilitate movement or provide resistance to movements, apply joint and soft tissue mobilisation or stretches, soft tissue and massage techniques.

The use of exercise programmes and providing advice on exercise to support self-management has historically been seen as integral to managing spondyloarthritis. With the localised and global effects on joints and tendons, other complications and the consequences for mobility, strength, function and fitness, determining the effectiveness of different exercise interventions for spondyloarthritis is important.

Hydrotherapy is another exercise intervention that has been traditionally seen as having a beneficial role in managing inflammatory disease. There are a number of water-based approaches that are encompassed within the term hydrotherapy, including water-based exercise that uses water to support the exercise process and includes a warmer water environment, exercising in non-hydrotherapy pools, and spa therapy and balneotherapy that may also use water with mineral salts or waterjets.

It is important to determine the effectiveness of different exercise interventions. This includes developing an understanding of how different factors affect outcomes. For example, structured or unstructured programmes, generic or individually tailored programmes, the type of exercise, the duration, intensity, frequency and mode of delivery, the degree of supervision, and the person's engagement and adherence with the exercise programme. It is also important to determine the level of knowledge and skill that is needed to support the exercise programme, and to explore the role of exercise in the short and longer term management of spondyloarthritis. Other considerations include whether there is evidence on the role and particular benefits of exercise at different stages of the disease or during different levels of disease activity, the impact on cardiovascular and non-articular complications or comorbidities, and when there are extra-articular factors to consider, such as skin or bowel involvement.

Acupuncture is becoming increasingly used in the management of health problems. Research on the benefits of acupuncture is growing, as are advances in the methodologies used to investigate its effectiveness and benefits. It is important to establish whether there is evidence of its effectiveness for use in the management of spondyloarthritis.

It is also important to establish the benefit of physical aids and appliances that may help relieve symptoms, prevent or reduce the loss of physical function, enable mobility and enhance a person's quality of life and ability to undertake their usual daily activities.

1 To date, research has been predominantly undertaken in axial spondyloarthritis, in particular
2 ankylosing spondylitis and often in established disease. There has been much less research
3 on the other disease groups and peripheral presentations. The lifelong, progressive nature of
4 spondyloarthritis means that it is challenging to evaluate the long-term benefits of
5 interventions such as exercise programmes. Interventions are commonly part of a multimodal
6 approach, and this creates challenges for determining the separate contributions of the
7 individual components.
8

6.1 Manual therapies for spondyloarthritis

Review question 14

- What is the effectiveness of manual therapies compared with standard care for managing spondyloarthritis?

6.1.1 Evidence review

The aim of this review was to assess the effectiveness of manual therapies for the management of symptoms of axial and peripheral spondyloarthritis.

The review focussed on identifying studies that fulfilled the conditions specified in Table 27.

Table 27: PICO inclusion criteria for the review question on manual therapies

Population	People diagnosed with spondyloarthritis

Interventions	• Manual therapies:
	  o soft tissue techniques (including massage, muscle energy technique and myofascial release)
	  o traction
	  o manipulation/mobilisation (including spinal manipulative therapy [SMT] and Maitland technique)
	  o mixed modality manual therapy (soft tissue techniques ± traction ± manipulation/mobilisation)
	• Minimum number or duration of treatment(s): 8 to 12 sessions or 3 months

Comparators	Standard care (usual care, treatment as usual, waiting list, delayed start of treatment, no treatment, placebo intervention)
Outcomes	Pain, adverse events, joint mobility, physical function, quality of life (condition specific measures preferred), imaging, composite measures (except BAS-G)

For full details of the review protocol, see Appendix C.

In total, 3,135 references were found for this review question and 5 studies were included; 1 randomised controlled trial (RCT; Widberg et al. 2009), 1 controlled clinical trial (CCT; Silva et al. 2012), 1 before-and-after study (Lubrano et al. 2006) and 2 cross-sectional studies (1 prospective; Lubrano et al. 2007 and 1 retrospective; Eppeland et al. 2013). A further study (Escalas et al. 2016) was identified during an update search.

6.1.1.1 Description of included studies

Details of the included studies are found in the evidence tables (see Appendix E), with GRADE profiles reported in Appendix G.

6.1.1.1.1 *Individualised (outpatient) programme of manual therapy and exercise vs. control (no treatment or standard care)*

Three studies evaluated the effects of individualised multimodal interventions comprising manual therapy and exercise on people diagnosed with ankylosing spondylitis (according to the modified New York criteria).

The first study, a 4-month RCT conducted in Sweden on 32 men (median age ranged from 35 and 36.5 years; median duration of disease ranged from 2.5 to 3.5 years; 75% taking NSAIDs, 44% taking DMARDs and 9% not on any medication) assessed the effects of an 8-week programme of self- and manual mobilisation and home exercises compared to no treatment (Widberg et al. 2009).

The second study, a 4-month CCT conducted in Brazil on 38 adults (26 men; mean age ranged from 35.3 to 44.3 years; mean duration of disease ranged from 7.1 to 10.1 years;

1 100% taking NSAIDs) assessed the effects of a 4-month global postural re-education
2 programme compared to usual care consisting of conventional stretching and breathing
3 exercises (Silva et al. 2012).

4 The third study, prospective cohort in France on 708 adults, assessed the impact on BASFI
5 scores of having versus not having physiotherapy (average of 23.1 sessions for those having
6 physiotherapy). (Escalas et al. 2016).

7 **6.1.1.1.2** *Group and individualised multimodal inpatient programme including manual therapy*
8 ***vs. no treatment***

9 Two studies (1 before-and-after; Lubrano et al. 2006 and 1 prospective case series; Lubrano
10 et al. 2007) from the same research group based in Italy evaluated the effects of a 3-week
11 multimodal inpatient programme comprising manual therapy (15 minutes of stretching with
12 neuromotor facilitation) at baseline and following treatment. A total of 71 adults (study sizes
13 19 and 52; 55 men and 16 women) with active ankylosing spondylitis (according to the
14 modified New York and ASAS criteria) were recruited (Lubrano et al. 2006 and 2007). The
15 mean (SD) ages were 41.3 (8.6) and 45.7 (10.0) years. The mean (SD) duration of disease
16 was 9.3 (6.0) and 7.8 (4.8) years. 84% and 81% were HLA-B27 positive. In both studies, all
17 participants were on NSAIDs, and a proportion was on DMARDs (53% and 33%), and in 1
18 study (Lubrano et al. 2006), some participants were on steroids (58%). Clinical peripheral
19 joint involvement, psoriasis and eye involvement were reported in 12, 6 and 9 participants
20 respectively.

21 The third study, a retrospective case series based in Norway analysed routinely collected
22 hospital data from 87 adults before and after participating in a 2-week inpatient rehabilitation
23 intervention consisting of group-based daily exercise (including water-based exercises) and
24 an individualised programme of massage, stretching, mobilisation/articulation and postural
25 advice, delivered by a multidisciplinary team comprising a rheumatologist, physiotherapist,
26 occupational therapist, social worker and secretary. Participants attending the programme
27 between January 2007 and June 2011 were diagnosed with axial spondyloarthritis
28 (according to the ASAS diagnostic criteria) and had imaging (X-ray, CT and/or MRI)-
29 confirmed sacroiliitis; 74% were diagnosed with ankylosing spondylitis according to the
30 modified New York criteria. Baseline study characteristics of the included population were: 60
31 men; 92.5% were HLA-B27 positive; mean (SD) age of 49.2 (10.0) years; mean (SD)
32 duration of disease of 14.4 (11.9) years; 62.1% were taking NSAIDs and 17.2% were taking
33 anti-TNFs (Eppeland et al. 2013).

34 The GRADE tables and forest plots are located in Appendix G.

35 **6.1.1.2** **Variations from protocol**

36 No relevant studies that delivered manual therapies as a single modality were identified.
37 However, several studies that administered manual therapies as part of a rehabilitation
38 programme consisting of exercise ± education/advice were identified and have been
39 reviewed for this question.

40 Comparative and observational studies that examined the effectiveness of manual therapies
41 in people diagnosed with spondyloarthritis were included. Papers were excluded if they:
42 • were guidelines, narrative reviews, case reports, case series with less than 10 people,
43   commentaries and editorials

44 • were not published in English language

45 • investigated a multimodal rehabilitation programme including a manual therapy
46   component that was not consistently administered, for example, a small proportion of
47   individuals may have received massage if indicated

48 • investigated a multimodal rehabilitation programme that included a manual therapy
49   component but the different components of the programme were variably administered,

1   for example, pain relieving local management (including local heat), water pool, group
2   gymnastics/exercise/keep fit club, individual gymnastics/exercise, individual physiotherapy
3   in suspension, stretching or mobilisation and massage selectively applied and/or
4   administered for different durations.

5   For the full list of excluded studies see Appendix F.

6.1.1.3 Minimal clinically important differences

7   A search in relation to axial spondyloarthritis, psoriatic arthritis and reactive arthritis of the
8   Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted
9   minimum clinically important difference thresholds for the outcomes in this review. It was felt
10   important to be consistent with the decisions made around MCIDs for the other treatment
11   questions which used pain as a primary outcome measure. Further, it was agreed to that any
12   consistently measurable reduction in pain (i.e. one that could be shown to be a significantly
13   greater reduction than random fluctuation) would be likely to be meaningful to patients, and
14   therefore statistically significant differences in pain outcomes were agreed to be clinically
15   meaningful. For other outcomes, where applicable, the GRADE default MID interval for
16   dichotomous outcomes of (0.8 to 1.25) was used.

6.1.2 Health economic evidence

18   A literature search was conducted jointly for all review questions in this guideline by applying
19   standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
20   total of 9,970 references was retrieved, of which none were retained for this review question.
21   Health economic modelling was not prioritised for this review question.

6.1.3 Evidence statements

6.1.3.1 Individualised programme of manual therapy and exercise vs. control (no treatment or usual care)

25   Low quality evidence from a randomised controlled trial (n=32) found that compared to no
26   treatment, men with ankylosing spondylitis receiving an 8-week individualised programme of
27   self- and manual mobilisation and home exercises showed a significant improvement in total
28   BASMI scores, but no differences in total BASFI and BASDAI scores immediately post-
29   treatment.

30   Very low quality evidence from a non-randomised controlled trial (n=35) found that compared
31   to conventional group-based stretching and breathing exercises, adults with ankylosing
32   spondylitis receiving a 4-month individualised global postural re-education programme
33   showed a significant improvement in joint mobility (measured using finger to floor distance
34   and cervical rotation), quality of life (measured using SF-36 physical and emotional
35   component scores), overall functional capacity (measured using Health Assessment
36   Questionnaire – Spondyloarthropathies) and total BASDAI scores. No significant between
37   group differences were observed in the Modified Schober Index.

38   Very low-quality evidence from 1 propensity matched cohort study of 689 people could not
39   differentiate proportions of people improving in BASFI scores at 6, 12 and 24 months
40   between people having and not having physiotherapy during the first 6 months of the study

6.1.3.2 Group and individualised multimodal inpatient programme including manual therapy vs. no treatment

43   Very low quality evidence from 2 studies from the same research group (n=71) found that
44   compared to pre-treatment scores, adults with active ankylosing spondylitis receiving a 3-
45   week intensive inpatient rehabilitation programme showed a significant improvement in pain

1 (measured using visual analogue scale), joint mobility (measured using modified Schober's
2 test and tragus to wall distance), quality of life (measured using EQ-5D VAS), total BASFI
3 scores and overall scores for the Leeds Revised Disability Questionnaire immediately post-
4 treatment. Follow-up assessments of 1 study (n=52) at 6 and 12 weeks that additionally
5 administered home exercises found significant improvement in pain, joint mobility, total
6 BASFI and Leeds Revised Disability Questionnaire scores. No significant improvement in
7 total BASDAI scores was found immediately post-treatment.

8 Very-low quality evidence from a retrospective case series (n=87) found that compared to
9 pre-treatment scores, adults with axial spondyloarthritis receiving a 2-week inpatient
10 rehabilitation programme showed a significant improvement in joint mobility (measured using
11 finger to floor distance), total BASFI, BASDAI and BASMI scores immediately post-treatment.
12 At a mean of 9.3 months follow-up, further assessment of the 3 BAS scales showed no
13 differences between the pre- and follow-up scores.

14 **6.1.4   Evidence to recommendations**

Relative value of different outcomes	The GDG agreed that BASFI and BASMI should be the primary outcomes considered when reviewing the effectiveness of these interventions. Pain and BASDAI were also considered important, but the GDG agreed that it was unlikely that these interventions would have a large effect on BASDAI scores.
Trade-off between benefits and harms	The GDG noted that there were large clinically important improvements in some of the reported outcomes immediately post-treatment for most of the included studies (for example, BASMI in men, joint mobility, quality of life, overall functional capacity, and total BASDAI score), but that the evidence indicated that these benefits decreased over time, and were not sustained at medium to longer term follow-up (up to a mean of 9.3 months). The GDG noted the lack of reported data on adverse events in all of the included studies. The GDG discussed the increased risk of fractures in people with ankylosed spines, particularly where osteoporosis is present and issues surrounding the use of high velocity low amplitude Grade 5 manipulative techniques. The GDG noted that none of the included studies described the use of Grade 5 manipulation. Although no evidence on high velocity low amplitude Grade 5 manipulation was available, given the nature of the condition, the GDG felt strongly that caution should be exercised and Grade 5 manipulative techniques should be avoided.
Consideration of health benefits and resource use	No health economic evidence was identified for this review question and this question was not prioritised for health economic modelling. The GDG agreed that, since it had not made a recommendation, it was highly unlikely any substantial increase in resource use would result.
Quality of evidence	The GDG noted that no evidence was found evaluating the effects of manual therapies administered as a single modality. The GDG highlighted that this was unsurprising given that in routine clinical practice, manual therapy would typically be offered as part of a package of care including exercises. The evidence review was therefore extended to include studies where manual therapy had been consistently applied to all (intervention) participants as part of a wider multi-modal rehabilitation programme. These studies were downgraded for indirectness in the GRADE assessment.

	The GDG agreed that the studies were of low to very low quality and that the manual therapy techniques used were variable. It was noted that overall, the study populations comprised mainly of people with ankylosing spondylitis. The GDG highlighted the differences in disease activity between the studies and noted that those studies which included individuals with active disease were likely to show greater improvements in outcomes than individuals who were considered stable. The GDG agreed that improvements were also reported to a greater extent in the non-randomised observational studies. The GDG agreed that it was not possible to identify the therapeutic contribution of the individual manual therapy components within the investigated complex interventions, and therefore felt that it would not be appropriate to make a specific recommendation on manual therapies alone.
Other considerations	The GDG noted that in-patient rehabilitation programmes are not widely available in the UK and therefore the evidence from the 3 observational studies in this setting is limited for making recommendations. The GDG noted the overlap of evidence in the studies of multimodal programmes with the guideline's physical intervention review questions on exercise and hydrotherapy. The GDG agreed that the presented evidence base for manual therapy would be reconsidered as appropriate (as multimodal intervention) for the review question on exercise.

6.1.5 Recommendations

No recommendation was drafted.

6.1.6 Research recommendations

7. What is the long-term effectiveness and cost-effectiveness of manual therapy as an intervention (without other concurrent physiotherapy) for both axial and peripheral spondyloarthritis, and does this effectiveness and cost-effectiveness change in different settings or between different delivery strategies?

Why this is important

Only a limited amount of low-quality evidence was found looking at the effectiveness of manual therapies, and therefore it was not felt possible to make any recommendations. Well conducted randomised controlled trials (in both axial and peripheral spondyloarthritis) comparing manual therapy interventions plus standard care to standard care alone would fill an important gap in the evidence base around which interventions provide effective symptom relief for people with spondyloarthritis.

6.2 Exercise for spondyloarthritis

Review Question 15

- What is the effectiveness of structured exercise compared with standard care for managing spondyloarthritis?

6.2.1 Evidence review

The aim of this review was to assess the effectiveness of structured exercise for the management of symptoms (axial and peripheral) of spondyloarthritis. The review focussed on identifying studies that fulfilled the conditions specified in Table 28.

Table 28: PICO table for question 15: exercise

Population	Patients with spondyloarthritis
Interventions	Structured exercise (individual, group, home, hospital, symptom/disease-specific)
Comparators	Standard care (including usual care, [treatment as usual], waiting list, delayed start of treatment, no treatment, placebo exercise)
Outcomes	Pain, adverse events, joint mobility, physical function, quality of life, imaging, composite measures

For full details of the review protocol please see Appendix C.

Randomised controlled trials were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well.

A systematic search and a hand search of the reference lists of systematic reviews identified 522 references. The references were screened on their titles and abstracts and 54 references were ordered for full text of which 53 were available.

Overall, 43 studies were excluded as they did not meet the eligibility criteria such as inappropriate study design (e.g. intervention study without a randomised control group) or non-primary studies (e.g. systematic review, editorial) A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F.

In total 10 articles were included, of which 1 (Karapolat 2009) featured evaluations of 2 different eligible exercise interventions compared with standard care. The studies fell into 5 different groups of comparisons:

- Unsupervised structured home exercise vs standard care
- Supervised individual structured exercise (outpatient) vs standard care
- Supervised individual structured exercise (inpatient) vs standard care
- Supervised structured group exercise vs unsupervised structured home exercise
- Supervised structured group exercise vs standard care

During an update search at the end of the guideline development process, a further 4 studies were found: three RCTs (Fang et al., 2016, Hseih et al., 2014, Jennings et al., 2015) and one quasi-RCT (Maseiro et al., 2014)

6.2.1.1 Description of included studies

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G.

Unsupervised structured home exercise vs standard care

Six RCTs (Fang et al., 2016, Hseih et al., 2014, Jennings et al., 2015, Kraag et al., 1990, Rodriguez-Lozano et al., 2013, Sweeney et al., 2002) including 1009 people with ankylosing spondylitis were identified for this comparison. One study used New York criteria (Kraag 1990), 4 used modified New York criteria (Fang 2016, Hseih 2014, Jennings 2015, Rodriguez-Lozano 2013), and the third did not state the diagnostic criteria (Sweeney 2002).

Structured exercise consisted of 14 weeks of physiotherapist-led exercise sessions (for up to 16 weeks) and a daily self-administered exercise programme.(Kraag 1990), a 2 hour face to face group information session about their condition, instructions to carry out a structured exercise programme at home of 6 months.(Rodriguez-Lozano 2013) and an exercise/educational video, containing an exercise regime suitable for all degrees of disease severity together with an educational booklet and a progress wall chart with reminder stickers. The study period was 6 months (Sweeney 2002). One study (Fang 2016) prescribed flexibility home based exercises of at least 3 60-minute sessions a week for 6 months, with fortnightly telephone follow up and monthly hospital appointments with a physiotherapist. In another (Hseih 2014), the intervention group undertook range of motion, strengthening and aerobic exercises following instruction. The final RCT intervention (Jennings 2015) involved aerobic exercise and stretching exercises, 3 times a week for 12 weeks. Comparison groups ranged from standard care alone, to advice on exercise without individual exercise therapy, to a more limited set of exercises than the intervention group.

Supervised individual structured exercise (outpatient) vs standard care

Two RCTs (Ince et al., 2006, Karapolat et al., 2009) including 75 people (67.2% male) with ankylosing spondylitis were included. Both studies diagnosed people according to modified New York criteria. The mean age of participants ranged from 33.4 years (SD=5.2) to 50.2 years (SD 12.4) and duration of disease ranged from 8.2 years (SD=5.7) to 20.6 years (SD 10.1). All participants in one study (Ince et al., 2006) received NSAIDs and sulfasalazine and in the second study (Karapolat et al., 2009) 40.5% were taking methotrexate, sulfasalazine or both.

Exercise interventions consisted of information, supervised exercise training, in the form of a multimodal exercise programme of three 50-minute sessions a week for 3 months (Ince 2006) and supervised walking or swimming (30 mins three times per week for six weeks alongside standard care of structured conventional exercise for 30 minutes once a day for six days.

Supervised individual structured exercise (inpatient) vs standard care

Two RCTs (Bulstrode et al., 1987, Kjeken et al., 2013) including 139 people (65.3% male in one study and gender was not specified in the second study) with ankylosing spondylitis were included. One study (Kjeken et al., 2013) included participants on the basis of modified New York criteria; the other did not state method of diagnosis (Bulstrode et al., 1987). The mean age was 49.9 years (no SD) and duration of disease was 15.5 years (no SD) in one study and was not reported in one study (Bulstrode et al., 1987). In one study (Kjeken et al., 2013), 75.8% of people received NSAIDs, 4.2% received standard DMARDs and 4.2% were taking biological DMARDs.

The exercise intervention consisted of admission for a 15-day in-hospital rehabilitation programme. involving daily passive stretching movements (Bulstrode et al., 1987) and a three-week inpatient rehabilitation programme of individualised treatment plan including exercise programme comprising multiple session of pool-based, gym-based and outdoor activities each week (Kjeken 2013). The control group in both studies consisted of standard care.

1 **Supervised structured group exercise vs unsupervised structured home exercise**

2 Two RCTs (Analay et al., 2003 and Cagliyan et al., 2007) including 97 people (83.5% were
3 male) with ankylosing spondylitis were included in this review. One study (Cagliyan et al.,
4 2007) used the modified New York diagnostic criteria while the other (Analay et al., 2003)
5 used the Amor criteria. The mean age of participants ranged from 34.3 years (SD 7.9) to
6 37.6 years (SD 11.3) and duration of disease was not reported in either study. One study
7 excluded people who were receiving DMARDs (Analay 2003)

8 The exercise intervention consisted of an education programme about the condition and a
9 six-week group exercise programme under the supervision of a physiotherapist (Analay et
10 al., 2003) and an education session about their condition with 3 months of physiotherapist-
11 supervised exercise at the hospital, with two 1-hour sessions per week (Cagliyan et al.,
12 2007). The control groups in both studies received standard care.

13 **Supervised structured group exercise vs standard care**

14 One RCT was identified for this comparison (Altan et al., 2012). The study population
15 included 30 men and 25 women, all with diagnosed ankylosing spondylitis according to
16 modified New York criteria. People with active peripheral symptoms were excluded. Mean
17 duration of disease across the two groups was 8.8 years (range 2-22 years). Mean age was
18 46.5 (SD=11.2) in the Pilates group and 43.6 (SD=10.1) in the control group. Participants
19 were allowed to continue previous medication but were requested not to use supplementary
20 drugs or change the usual dosages throughout the study period and were asked not to take
21 any pain killers in the morning of the assessment day. 31% were regularly taking NSAIDs,
22 32% sulfasalazine, 21% biological DMARDs and 17% took no regular medication.

23 The intervention was a Pilates-based exercise programme of 1 hour, 3 times a week, led by
24 a certified trainer (30 participants). The control group (25 participants) received usual care,
25 and were instructed to continue participating in their usual physical activity.

26 One quasi-randomised controlled trial was also identified (Maseiro et al., 2014). Sixty nine
27 people aged 18 – 65 with ankylosing spondylitis whose condition had been stabilised with
28 anti-TNF inhibitors were recruited, of whom 64 were included in the analysis. The
29 intervention group received an education component developed by an interdisciplinary team,
30 comprising 2 meetings of 3 hours each in groups of 8-12 people. They also received an
31 exercise intervention involving 12 twice-weekly sessions of 60 minutes each, delivered by a
32 physiotherapist in a group setting (6-8 patients). Following the group sessions, participants
33 were encouraged to continue exercise at home. A second group only received the education
34 intervention, and a third group received no intervention but continued on standard biological
35 DMARDs.

36 **6.2.1.2 Minimal clinically important differences**

37 A search in relation to axial spondyloarthritis, psoriatic arthritis and reactive arthritis of the
38 Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted
39 minimum clinically important difference thresholds for the outcomes in this review. It was felt
40 important to be consistent with the decisions made around MCIDs for the other treatment
41 questions which used pain as a primary outcome measure. Further, it was agreed to that any
42 consistently measurable reduction in pain (i.e. one that could be shown to be a significantly
43 greater reduction than random fluctuation) would be likely to be meaningful to patients, and
44 therefore statistically significant differences in pain outcomes were agreed to be clinically
45 meaningful. For other outcomes, where applicable, the GRADE default MID interval for
46 dichotomous outcomes of (0.8 to 1.25) was used.

1 ## 6.2.2 Health economic evidence

2 A literature search was conducted jointly for all review questions in this guideline by applying
3 standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
4 total of 9,970 references was retrieved, of which none were retained for this review question.
5 Health economic modelling was not prioritised for this review question.

6 ## 6.2.3 Evidence statements

7 **Unsupervised structured home exercise vs standard care**

8 High quality evidence found that participants receiving unsupervised structured home
9 exercise had improved quality of life (1 RCT, n=756), reduced finger to floor distance (1 RCT,
10 n=48), and improved BASFI score (5 RCTs, n=1034) when compared with standard care.

11 Low to moderate quality evidence found no significant difference in BASDAI score (5 RCTs,
12 n=1034) BASMI score (2 studies, n=104) or BASG score (2 RCTs, n=174)

13 Very low quality evidence found no significant difference in pain (3 RCTs, n= 959) in
14 participants receiving home exercise compared with standard care.

15 **Supervised individual structured exercise (outpatient) vs standard care**

16 Very low quality evidence found no significant difference in finger-floor distance (2 RCTs,
17 n=80), BASMI score (2 RCTs, n=68) or pain (2 RCTS, n=38) in participants receiving
18 structured supervised outpatient exercise compared with standard.

19 **Supervised individual structured exercise (inpatient) vs standard care**

20 Moderate quality evidence from one RCT (n= 95) found no significant difference in BASDAI,
21 BASMI, and BASFI scores in participants receiving inpatient structured exercise compared
22 with standard care.

23 **Supervised structured group exercise vs unsupervised structured home exercise**

24 Low to moderate quality evidence found no significant difference in BASFI score or stiffness
25 (1 RCT, n=45), finger-floor distance, or pain (2 RCTs, n = 91) in those receiving structured
26 supervised group exercise compared with structured unsupervised home exercise.

27 **Supervised structured group exercise vs standard care**

28 Low to moderate quality evidence (2 RCTs, n=97) found an improvement in BASDAI score
29 and BASFI score in those receiving structured supervised group exercise compared with
30 standard care.

31 Very low to moderate quality evidence (2 RCTs, n=97) found no significant difference in
32 BASMI score or quality of life in those receiving structured supervised group exercise
33 compared with standard care.

34 ## 6.2.4 Evidence to recommendations

Relative value of different outcomes	The GDG discussed which of the available outcome measures were important for assessing the possible benefits of exercise-based interventions and noted the following: - In terms of function, due to the progressive nature of the condition, healthcare professionals may not be looking for improvement over long term but rather a 'no worsening' of symptoms. - BASMI is a composite measure used routinely in clinical practice,

	the individual aspects of which are reported separately in some studies
- Schober test + intermalleolar distance (from BASMI) /finger to floor can be pooled if necessary, and is used in clinical practice
- Pain at rest/activity may not show as much benefit as other outcomes
- Fatigue is an important outcome
- Stiffness
- Joint mobility: lumbar flexibility is useful depending on state and site of the of the disease; cervical flexion can be combined with thoracic;
- Chest expansion is part of modified New York criteria. Although this is not very reliable it can be useful in clinical practice depending on stage of disease.
Trade-off between benefits and harms	The GDG did not expect structured exercise to have a detrimental effect on the disease process (inflammation).
The GDG would expect to see improvements over the short term, however they were looking more for longer term benefits such as maintaining mobility and function.
There is large variation in practice across the NHS so any recommendations will have an impact on patients nationwide
No evidence for peripheral symptoms was identified; however the GDG agreed that exercise would be beneficial for people with spondyloarthritis.
Economic considerations	No health economic evidence was found and this review question was not prioritised for health economic modelling.
Whilst economic modelling was not undertaken for this question, the evidence assembled for the economic model addressing the referral/diagnosis questions showed that improving functional ability, as measured by BASFI, is a critical determinant of both patients' quality of life and background health and social care costs (see Appendix H). The GDG noted that evidence of benefit for physiotherapy was strongest, in this domain. The analysis also considered the costs of specialist physiotherapy as a parameter in the simulation of downstream consequences of diagnosis. The evidence identified here suggested that a course of physiotherapy costs a little over £200 per person. Taking these 2 pieces of evidence from the diagnosis model together, the GDG was confident that specialist physiotherapy is likely to be cost effective.
Quality of evidence	All evidence was for axial symptoms in ankylosing spondylitis though it was not always clear which diagnostic criteria was used.
Overall there was agreement within the GDG that the evidence was low quality due to the following factors:
- Few people on NSAIDs included in the studies
- The long disease duration experienced by people in many of the included studies could mean that it would be hard to detect a benefit of exercise, particularly if disease has progressed to a stage where irreversible changes in function have occurred.
- Self-selecting populations: may have influenced the effects of standard care in these population (narrowing gap between standard care and interventional group)
- Under-reporting of baseline characteristics raises problems with interpreting outcomes
- The duration of follow up may be too short to allow the full range of benefits to be recorded and a for a positive response to be maintained
The intervention in the study examining Pilates was not considered to be a standard Pilates exercise session, but was modified to suit this group of patients. The GDG considered that the intervention as

	described was akin to an intervention they would recommend.
	There was evidence of benefit for both a structured exercise programme following physiotherapist-led induction session and also for supervised and structured stretching exercise (as a component of Pilates) led by a physical trainer. The GDG noted the common components of these interventions in terms of stretching, breathing, spinal extension and motion exercises and therefore opted to recommend a physiotherapist-led approach.
	The GDG discussed the role of the physiotherapist in some interventions and believed that this should be a specialist physiotherapist.
	The GDG also agreed that any structured exercise programmes should be introduced by a physiotherapist to support and demonstrate the more difficult exercises.
Other considerations	The GDG noted that the number of face to face sessions of physiotherapy-led exercise is often limited, so a recommendation in this area would help resource these services. The GDG acknowledged that where variation in access to physiotherapy-led exercise currently exists, this would have a varying impact on resources, with the greatest impact felt where people with spondyloarthritis are not currently receiving guided exercise programmes.
	The evidence presented was limited to studies on people with ankylosing spondylitis and the GDG believed that to restrict the recommendations to those with diagnosed ankylosing spondylitis rather than more broadly applying them to axial spondyloarthritis would mean that women in particular might not benefit as historically there has been under-diagnosis of ankylosing spondylitis in women.
	The GDG considered that the findings could be extrapolated from ankylosing spondylitis to axial symptoms in other spondyloarthritis conditions. They acknowledged, however, that there was no evidence presented on the evaluation of exercise for managing peripheral spondyloarthritis symptoms, and did not opt to make a recommendation on exercise in peripheral joints.

6.2.5 Recommendations

17. Refer people with axial spondyloarthritis to a specialist physiotherapist to start a structured exercise programme, which should include:

- stretching
- deep breathing
- spinal extension
- range of motion exercises for the lumbar, thoracic and cervical sections of the spine

18. aerobic exercise.

-

6.2.6 Research recommendations

8. What is the short- and long-term effectiveness and cost-effectiveness of structured exercise programs for peripheral spondyloarthritis, and does this effectiveness and cost-effectiveness change in different settings or between different delivery strategies

1 **Why this is important**

2 Whilst moderate-quality evidence was found looking at the effectiveness of structured
3 exercise programs in axial spondyloarthritis, no evidence was found for people with
4 peripheral spondyloarthritis, where it is also believed there may be a positive effect
5 (particularly in people with axial involvement). Well conducted long-term randomised
6 controlled trials comparing structured exercise programs interventions plus standard care to
7 standard care alone would fill an important gap in the evidence base around which
8 interventions provide effective symptom relief for people with peripheral spondyloarthritis.
9

6.3 Hydrotherapy for spondyloarthritis

Review Question 16

- What is the effectiveness of hydrotherapy compared with standard care for managing spondyloarthritis?

6.3.1 Evidence review

The aim of this review was to assess the effectiveness of hydrotherapy for the management of symptoms (axial and peripheral) of spondyloarthritis.

The review focused on identifying studies that fulfilled the conditions specified in Table 29.

Table 29: PICO table for question 16: hydrotherapy

Population	Patients with spondyloarthritis
Interventions	Hydrotherapy
Comparators	Standard care (including usual care, treatment as usual, waiting list, delayed start of treatment, no treatment, placebo intervention)
Outcomes	Pain, adverse events, joint/spinal mobility, physical function, quality of life, imaging, composite measures

For full details of the review protocol please see Appendix C.

Randomised controlled trials were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well. Observational studies were also sought, in anticipation that good quality RCTs were unlikely to be available.

A systematic search and a hand-search of the reference lists of systematic reviews identified 1,623 references. The references were screened on their titles and abstracts and 69 studies were ordered for full text of which 68 were available.

Fifty-six studies were excluded as they did not meet the eligibility criteria such as inappropriate study design (for example, case reports and case series with fewer than 10 cases) or non-primary studies (for example, editorials). A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F.

A total of 5 RCTs and 7 observational studies was included.

6.3.1.1 Variations from protocol

Randomised controlled trials of therapist-led hydrotherapy were rare, so studies of balneotherapy and spa therapy (where the therapy explicitly involved a water-based therapy such as heated baths) were also included.

6.3.1.2 Description of included studies

The included RCTs covered the following comparisons:

- Active hydrotherapy (i.e. movement in water) vs standard care in people with axial symptoms
- Passive hydrotherapy (i.e. bathing) vs standard care in people with axial symptoms
- Passive hydrotherapy (i.e. bathing) with electrical current vs standard care in people with either axial or peripheral symptoms or both.

The included observational studies covered the following:

- Active hydrotherapy alone

1 • Passive hydrotherapy alone

2 • Active hydrotherapy as part of a complex physical therapy intervention.

3 **6.3.1.3 Description of included studies**

4 Evidence tables for included studies can be found in Appendix E, with GRADE profiles
5 reported in Appendix G.

6 *6.3.1.3.1 RCT evidence*

7 **Active hydrotherapy vs standard care in people with axial symptoms**

8 There was 1 RCT of active hydrotherapy vs standard care (Ciprian et al., 2013). Most
9 intervention group participants (n=15) were male (93.3%), 86.7% were HLA-B27 positive and
10 73.3% were treated with etanercept. A further 26.6% were treated with infliximab. Their mean
11 age was 47.8 (SD =10) and disease duration was 13.9 years (SD=8.6). In the control group
12 (n=15) 93.3% were men, 93.3% were HLA-B27 positive, 66.7% were on etanercept and
13 33.3% were on infliximab. Mean age was 45.6 (SD=11.8) and mean disease duration was
14 13.2 years (SD=8.8).

15 Participants in the intervention group received 10 sessions of spa therapy over a 10-week
16 period. The therapy involved application of a heated (40–55°C) mud pack to the entire spinal
17 area for 15 minutes followed by immersion to neck level in a thermal bath tank (37–38°C) for
18 10 minutes. Participants then had a group rehabilitation session in a pool of thermal water
19 (32–34°C) where they performed exercises (spine mobilisation, muscular spine
20 strengthening, respiratory kinesitherapy) under the supervision of a physiotherapist. The
21 thermal water contained mineral salts and had been obtained from a well. Participants in the
22 control group are not reported to have received any non-pharmacological intervention.

23 **Passive hydrotherapy vs standard care in people with axial symptoms**

24 Three RCTs of passive hydrotherapy interventions were identified (Altan et al., 2006, Cozzi
25 et al., 2007 and Yurtkuran et al., 2005). Two RCTs (Altan, Yurtkuran) of 95 people studied
26 patients with ankylosing spondylitis, all diagnosed using New York or modified New York
27 criteria. One study (Cozzi) involved 24 people with spondylitis and inflammatory bowel
28 disease. All fulfilled ESSG criteria and had IBD (Crohn's disease or ulcerative colitis) that had
29 been diagnosed by clinical, endoscopic, histological and radiological criteria. None had
30 peripheral disease. One study (Altan) did not report patient characteristics or disease
31 duration at baseline; of the other 2, mean age ranged from 41.4 (SD=11.8) to 57 (SD=7), and
32 mean duration of spondylitis symptoms ranged from 6.8 (SD=6.5) to 12 (SD=5) years. In
33 1 study (Yurtkuran), all participants received NSAIDs during the intervention period; in
34 another (Altan) all were allowed to continue their regular medication but were asked to make
35 no changes, and in the third (Cozzi) no NSAIDs or corticosteroids were permitted
36 (paracetamol was allowed as needed).

37 Interventions were mud packs (42–45°C, 15 mins) followed by thermal mineral baths (37–
38 38°C, 10 mins) for 12 sessions over 2 weeks (Cozzi) compared with standard care;
39 balneotherapy (39°C spa water, 30 mins/day for 3 weeks) followed by 2 hours of bed-rest
40 plus an instructed home exercise programme (30 mins/day for 24 weeks) comprising
41 respiration-postural exercises and dorsal/lumbar extension exercises (Altan), compared with
42 the exercise programme alone; and balneotherapy (37°C spring-water, 20 mins/day,
43 5 days/week for 3 weeks) plus 1000 mg naproxen and 400 mcg misoprostol/day (Yurtkuran),
44 compared with NSAIDs alone.

1 **Passive hydrotherapy with electrical current vs standard care in people with axial**
2 **and/or peripheral symptoms**

3 One study (Gurcay et al., 2008) described an RCT of Stanger bath therapy in which an
4 electrical current is applied to a bath of warmed water. The 29 participants in the intervention
5 group had a mean age of 40.2 (SD=10.38) and a mean disease duration of 16.21 years
6 (SD=10.22). 6.9% were female, 20.1% had peripheral symptoms, 51.7% axial and 27.6%
7 had both. In the control group (n=28) the mean age was 41.3 (SD =8.59) and mean disease
8 duration was 13.53 years (SD=9.33). 21.4% of control group participants were female, 7.1%
9 of the group had peripheral symptoms, 60.7% had axial symptoms and 32.1% had both.

10 The intervention group had 15 sessions over 3 weeks of exercise plus bath therapy. The
11 exercise programme was taught to participants individually by a physiotherapist, and then
12 carried out without supervision at home. It included range of motion, muscle strengthening,
13 respiration and postural exercises. The bath therapy involved bathing in warmed tap water
14 (36–37°C) in an adapted bath to which diadynamic (DD) current was applied at varying
15 intensity. Participants in the control group received the same exercise programme but no
16 bath therapy.

17 **6.3.1.3.2** *Observational evidence*

18 **Active hydrotherapy alone in people with axial or axial and peripheral symptoms**

19 One study (Robertson et al., 2004) was a retrospective cohort study in which UK ankylosing
20 spondylitis patients diagnosed with modified New York criteria at the Royal Cornwall Hospital
21 received annual questionnaires over a period of 5 years.

22 Seventy four people provided data for at least 3 years (range 3–5 years follow up), of whom
23 24.3% were female, the mean age was 48.5 years (SD 11.24) and the mean disease
24 duration was 21.1 years (SD 10.63). Peripheral symptoms were present in 52.7% of
25 participants. 5.4% of people received DMARDs and 78.3% took regular NSAIDs. Of the 74
26 people in the cohort, 17 reported receiving regular hydrotherapy.

27 **Passive hydrotherapy alone in people with axial symptoms**

28 Two non-comparative cohort studies were identified of passive hydrotherapy without other
29 physical therapy interventions in people with ankylosing spondylitis. One study (Annegret et
30 al., 2013) was the control group (n=19) of a randomised controlled trial of radon therapy
31 baths in a group which also included patients with rheumatoid arthritis and osteoarthritis. The
32 other (Tishler et al., 1995) was a non-controlled randomised intervention of 14 people
33 (participants were randomly selected from a hospital list). Mean age ranged from 45.3 (range
34 33–65) to 59.6 (SD=12.9), mean disease duration was reported in 1 study (Tishler) as
35 13.6 years (range 2–18) and 18.1% of participants were female. One study (Tishler) reported
36 that participants were diagnosed according to modified New York criteria, 10 were using only
37 NSAIDs, 2 NSAIDs and analgesics and 2 using only analgesics.

38 Both studies involved a period of regular bathing (frequency daily to 3 days) in warm water
39 (36–38°C) for 20–30 minutes over a period of 2–4 weeks. One study (Tishler) additionally
40 applied daily mud packs to the lower back for 20 minutes at an initial temperature of 45°C.
41 Follow up period ranged from 12 weeks to 9 months.

42 **Active hydrotherapy as part of a complex physical therapy intervention in people with**
43 **axial or axial and peripheral symptoms**

44 Four studies presented complex interventions of physical therapy including hydrotherapy
45 (Aydemir et al., 2010; Colina et al., 2009; Eppeland et al., 2013; van Tubergen et al., 2001).
46 Two (Aydemir, Colina) were prospective non-randomised interventional studies, one

(Eppeland) was a retrospective case series and one presented the outcomes of the control group of a prospective RCT (van Tubergen).

People in the 3 prospective studies (Aydemir, Colina, van Tubergen; n=97) had mean ages ranging from 24.39 (SD=2.97) to 48 (SD=10) and a mean disease duration ranging from 4.71 (SD=1.86) to 10 (SD=6) years. All 3 studies were of people with ankylosing spondylitis, with 2 reporting diagnosis by modified New York criteria and 1 (Colina) not reporting the criteria used. The latter contained 7 participants with peripheral symptoms. In 2 studies (Aydemir, van Tubergen) the percentage of participants who were female was 10.4%. All participants in 1 study (Colina) received etanercept as part of the intervention. In 1 (Aydemir), 6 months of taking sulfasalazine and indomethacin was an inclusion criterion and in the third (van Tubergen) 12.8% took DMARDs and 92.3% took NSAIDs.

The retrospective case series (Eppeland), which examined records of 87 people with axial spondylitis, reported each outcome on different numbers of people. Overall their study population was 31.0% female, had a mean age of 49 (no SD) and a mean disease duration of 14 years (no SD). All patients fulfilled ASAS diagnostic criteria and had sacroiliitis confirmed by imaging. Additionally 92.5% were HLA-B27 positive, 62 were current users of NSAIDs and 17% took anti-TNFs.

All the interventions involved hydrotherapy alongside other physical therapies. Three involved intensive rehabilitation programmes for 5–7 days per week (Aydemir, Colina, Eppeland) over a period of 1–3 weeks with follow up time of 2 weeks to 6 months. Programmes typically involved water-based exercises alongside other physical therapies such as land-based exercises, stretching, postural exercises and respiratory exercises. One study (van Tubergen) featured a control group who received weekly physical therapy comprising 1 hour of physical exercises, 1 hour of hydrotherapy and 1 hour of sport, and were followed up over a period of 40 weeks.

6.3.1.4 Minimal clinically important differences

A search in relation to psoriatic arthritis and reactive arthritis of the Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted minimum clinically important difference thresholds for the outcomes in this review. It was felt important to be consistent with the decisions made around MCIDs for the other treatment questions which used pain as a primary outcome measure. Further, it was agreed to that any consistently measurable reduction in pain (i.e. one that could be shown to be a significantly greater reduction than random fluctuation) would be likely to be meaningful to patients, and therefore statistically significant differences in pain outcomes were agreed to be clinically meaningful. For other outcomes, where applicable, the GRADE default MID interval for dichotomous outcomes of (0.8 to 1.25) was used.

6.3.2 Health economics evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

6.3.3 Evidence statements

6.3.3.1 RCT evidence

Active hydrotherapy vs standard care in people with axial symptoms

Very low-quality evidence from a single RCT (n=30) found no significant difference in total BASMI score, total BASDAI score, total HAQ score or self-reported pain measured on a

visual analogue scale in participants receiving active hydrotherapy compared with standard care.

Passive hydrotherapy vs standard care in people with axial symptoms

Very low quality evidence found no significant difference in functional capacity (3 RCTs, n=172), BASDAI (2 RCTS, n=78), finger-floor distance (1 RCT, n=37), or self-reported pain measured on a visual analogue scale (3 RCTs, n=115) for participants receiving passive hydrotherapy compared with standard care.

Passive hydrotherapy with electrical current vs standard care in people with axial and peripheral disease symptoms

Very low-quality evidence from a single RCT (n=57) found that participants receiving passive hydrotherapy with electrical current showed an improvement in total BASMI score, total BASFI score, total BASDAI score and total ASQoL score, compared with participants who received standard care.

6.3.3.2 Observational evidence

Active hydrotherapy alone in people with axial symptoms

Very low-quality evidence from a single retrospective cohort study (n=74) in which some people reported receiving active hydrotherapy found no statistically significant change in total BASFI score, compared with a statistically significant decline in function in people who did not report receiving hydrotherapy.

Passive hydrotherapy alone in people with axial symptoms

Low-quality evidence from 1 non-controlled intervention study of passive hydrotherapy (n=14) found a reduction in morning stiffness time and a reduction in finger floor distance in people with axial symptoms.

Low-quality evidence from 2 non-controlled intervention studies of passive hydrotherapy (n=33) found no significant change in total BASFI score or self-assessed pain score for people with axial symptoms.

Active hydrotherapy as part of a complex intervention in people with axial symptoms

Very low-quality evidence from a non-randomised controlled intervention (n=30) of active hydrotherapy as part of a complex intervention found an improvement in quality of life as measured by the EQ-5D instrument in people with axial symptoms in the intervention group. Very low-quality evidence from a retrospective case series found an improvement in total BASMI score (n=87), total BASDAI score (n=59) and finger-floor distance (n=49). Both studies found very low quality evidence of an improved BASFI score (n=87).

Very low-quality evidence from a non-randomised intervention study (n=28) presented ambiguous findings on total BASFI score (an increased score was detected, but there was no measure of error).

Very low-quality evidence from a non-randomised intervention study (n=28) of active hydrotherapy as part of a complex intervention found no statistically significant change in total BASMI score, or in pain or physical function (both as measured by SF-36 domains). Low-quality evidence from the control group of an RCT (n=39) found no statistically significant change in total BASFI score, quality of life measured by ASQoL and duration of morning stiffness. Low- to very low-quality evidence from both studies (n=67) did not find a significant change in total BASDAI score.

1 6.3.4 Evidence to recommendations

Relative value of different outcomes	The GDG highlighted pain and mobility/function (including composite measures) as the most important outcomes. This was followed by quality of life, finger-floor distance and morning stiffness. The GDG noted that although RCTs are the preferred study design, longer term follow-up might be needed to evaluate the benefits of hydrotherapy, which might only be available from observational studies.
Trade-off between benefits and harms	The GDG did not identify any specific harms associated with undergoing hydrotherapy in people for whom the intervention is appropriate. There was no particular evidence of adverse events in the studies that were reviewed. It was noted that there were some risks associated with excessive or incorrect forms of exercise in this population.
	The primary benefit, based on the GDG's experience, was considered to be a lack of progression of disease. The GDG's expectation was that hydrotherapy would relax muscles, enabling to a greater range of movement.
	It was noted that hydrotherapy will often be delivered in group settings, which may have pros and cons for individual members of the group.
Economic considerations	No economic evidence was identified for this question, and it was not prioritised for health economic modelling. The GDG noted that the main reason for hydrotherapy not being widely available was perceptions around its cost. It was noted that the key cost driver was the construction of the facilities in the first place, rather than the costs to use those facilities once they exist.
	Whilst economic modelling was not undertaken for this question, the costs of hydrotherapy were considered as a parameter in the economic model built to address the referral/diagnosis questions (see Appendix H). The evidence identified here demonstrates that the incremental cost of using hydrotherapy facilities, once they exist, is low. Therefore, while the GDG agreed it would not be appropriate to recommend an expansion in the number of hydrotherapy facilities available, it felt that the limited economic data suggested that any benefits – even if small – delivered by therapy in existing facilities would justify the small costs.
Quality of evidence	All of the evidence identified was of low to very low quality.
	Some studies were not directly representative of UK clinical practice, where people usually receive up to 1 session a week, usually in an outpatient setting.
	Many studies involved complex interventions/multi-modal approaches, which made it hard to identify the specific effects of hydrotherapy. Older studies (pre-1999) reported a different range of outcome measures, as this era pre-dates the availability of the Bath Indices; it is therefore difficult to directly compare studies pre- and post- this era.
	It was noted that the findings were stronger in the context of hydrotherapy as part of a complex intervention, rather than on its own.
	The GDG acknowledged that it is difficult to conduct studies to determine the effectiveness of hydrotherapy, due to the variation in access to suitable centres nationally, and the delivery of it alongside other forms of physiotherapy.
Other considerations	The GDG agreed that hydrotherapy may be of particular use in the management of flare episodes, but that the included studies did not focus on this aspect of the condition.
	The GDG noted that some studies included participants who had advanced disease, or recruited an older population rather than

people with new-onset disease. It was agreed that there would be less chance of observing a benefit in these populations. The GDG commented that people with less severe disease are more likely to be in work and less able to take time off to participate in the more intensive study regimes.

The GDG agreed that, particularly in advanced disease, an absence of deterioration should be considered a positive outcome measure, as it may be the case that in some people a measurable improvement is unlikely to be achievable at their stage of disease.

The GDG noted that the evidence identified did not reflect their personal experiences (as practitioners or people with spondyloarthritis) with respect to patient-reporting of benefits of hydrotherapy to GPs or patient groups, and the number of requests received for repeat referrals.

It was noted that there was little outcome data on peripheral symptoms, and none of the studies focussed on people with peripheral spondyloarthritis. The GDG also felt that hydrotherapy was likely to have greater benefits in people with axial spondyloarthritis.

The GDG opted to make a recommendation based on the evidence and their own clinical experience, giving particular weight to studies with long term follow up.

6.3.5 Recommendations

19. Consider hydrotherapy as an adjunctive therapy to manage pain and maintain or improve function for people with axial spondyloarthritis.

6.3.6 Research recommendations

9. What is the short- and long-term effectiveness and cost-effectiveness of hydrotherapy in improving patient-reported outcomes in spondyloarthritis, and does this effectiveness and cost-effectiveness differ between hydrotherapy in a hydro pool or a standard swimming pool?

Why this is important

Whilst evidence around hydrotherapy does exist in the form of short-term randomised controlled trials and longer-term observational studies, there is currently a lack of long-term randomised controlled trials which have been conducted looking at the effectiveness of hydrotherapy for people with spondyloarthritis. Well conducted long-term RCTs (in both axial and peripheral spondyloarthritis) comparing hydrotherapy plus standard care to standard care alone would fill an important gap in the evidence base around which interventions provide effective symptom relief for people with spondyloarthritis. Further, the majority of the concerns around the affordability of hydrotherapy as an intervention are based on it having to be conducted in a specialist hydrotherapy pool. It would therefore be important to know whether a much cheaper and more available hydrotherapy intervention, using a standard swimming pool, offers equivalent benefits.

10. What is the effectiveness and cost-effectiveness of hydrotherapy in managing flares in people with spondyloarthritis, and does this effectiveness and cost-effectiveness differ between hydrotherapy in a hydro pool or a standard swimming pool?

1   **Why this is important**

2   No evidence was identified about the benefits of hydrotherapy for managing flares, an
3   important gap in the evidence base as this is one of the situations it is felt likely to have the
4   greatest benefits. There is thus the need for randomised controlled trials (following people up
5   for at least the entire duration of their flare) comparing hydrotherapy plus standard care to
6   standard care alone. Further, the majority of the concerns around the affordability of
7   hydrotherapy as an intervention are based on it having to be conducted in a specialist
8   hydrotherapy pool. It would therefore be important to know whether a much cheaper and
9   more available hydrotherapy intervention, using a standard swimming pool, offers equivalent
10   benefits.
11

6.4 Acupuncture for spondyloarthritis

Review Question 17

- What is the effectiveness of acupuncture compared with sham acupuncture and standard care for managing spondyloarthritis?

6.4.1 Evidence review

The aim of this review question was to determine the clinical effectiveness of acupuncture in the management of spondyloarthritis symptoms (axial and peripheral).

The review focused on identifying studies that fulfilled the conditions specified in Table 30.

Table 30: PICO table for question 17: acupuncture

Population	People with spondyloarthritis
Interventions	Acupuncture
Comparators	Sham acupuncture (pre-specified) or standard care (post-hoc decision)
Outcomes	Pain, adverse events, joint/spinal mobility, physical function, quality of life, imaging, composite measures

For full details of the review protocol please see Appendix C.

Randomised controlled trials were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well. A total of 80 studies were identified for this question, of which 11 were ordered for full-text review. The two studies finally included in the review are described below, with reasons for the exclusion of the other 9 given in Appendix F.

6.4.1.1 Variations from protocol

Due to the limited number of eligible studies comparing acupuncture to sham acupuncture, RCTs which used standard care as a comparator were also included.

6.4.1.2 Description of included studies

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G.

Acupuncture vs sham acupuncture

One RCT (Emery and Lythgoe, 1986) randomised participants to either acupuncture or sham acupuncture. The study population included 10 people diagnosed with ankylosing spondylitis but the diagnostic criteria used were not reported. No baseline data on gender, mean age or duration of disease was reported. Six of the 10 participants were taking anti-inflammatory drugs, but it is not reported which group these were in.

Acupuncture vs standard care

One RCT (Jia et al., 2006) randomised participants to either moxibustion (the burning of dried mugwort on particular parts of the body) or acupuncture or a 'standard care' control group. Only data from the acupuncture and standard care groups is used in this review. The study population included 60 people (75% were male) diagnosed with ankylosing spondylitis but the diagnostic criteria used were not reported. The mean age of the participants was 22.6 (SD=5.1) and 22.0 (SD=5.4) years, and the duration of symptoms was 4.3 (SD=5) and 4.4 (SD=3.2) years for the acupuncture and standard care groups respectively. All study participants received methotrexate and salicylazosulfapyridine.

6.4.1.3 Minimal clinically important differences

A search in relation to axial spondyloarthritis, psoriatic arthritis and reactive arthritis of the Core Outcome Measures in Effectiveness Trials (COMET) database did not yield accepted minimum clinically important difference thresholds for the outcomes in this review. It was felt important to be consistent with the decisions made around MCIDs for the other treatment questions which used pain as a primary outcome measure. Further, it was agreed to that any consistently measurable reduction in pain (i.e. one that could be shown to be a significantly greater reduction than random fluctuation) would be likely to be meaningful to patients, and therefore statistically significant differences in pain outcomes were agreed to be clinically meaningful. For other outcomes, where applicable, the GRADE default MID interval for dichotomous outcomes of (0.8 to 1.25) was used.

6.4.2 Health economics evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

6.4.3 Evidence statements

Acupuncture vs sham acupuncture

Very low-quality evidence from a single RCT (n=10) found no significant difference between acupuncture and sham acupuncture for pain and stiffness. No evidence was found for the remaining outcomes.

Acupuncture vs standard care

Moderate-quality evidence from a single RCT (n=60) found a significant benefit for acupuncture vs standard care for joint and spine mobility (as measured by the finger-floor distance).

Low-quality evidence form the same study found no difference between the 2 groups for the following outcomes: swollen and painful peripheral joints, morning stiffness.

6.4.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that the predominant outcome of interest is pain, as this is the main outcome where it is hypothesised there may be a benefit from acupuncture.
Trade-off between benefits and harms	Any general risks associated with acupuncture are likely to apply to this population. People may also incur personal costs if seeking out acupuncture in private practice.
	If people seek complementary therapies outside of standard rheumatology care, it may be difficult to monitor the effectiveness of these therapies if they are not recorded on medical records. It may also affect the adherence to interventions prescribed as part of standard care.
	If acupuncture is effective, it may be perceived by some patients to be a more acceptable intervention than pharmacological agents.
Economic considerations	No health economic evidence was found and this review question was not prioritised for health economic modelling. The relatively small number of people who would be eligible for acupuncture to treat spondyloarthritis makes it unlikely that services would be set up specifically for this group. Therefore, people are likely to be seen as part of general acupuncture services, and hence the additional cost

	of treating a person would be low.
Quality of evidence	In considering the included evidence the GDG noted the following: • The studies all had small sample sizes (one study had n=10), and therefore even if acupuncture had benefits it would be HARD TO detect them in these trials . • The lack of sham acupuncture in the larger study may inflate observed differences, as other trials have acupuncture have shown there is a large placebo effect when sham acupuncture is not used as the comparator. • Difficulties may arise in administering sham acupuncture, leading to a risk that participants are not properly blinded to the intervention, which may bias their perceptions of the efficacy. • The intervention in 1 study was a very intensive course of treatment with 1 session every other day for 6 months; this might not be achievable in a non-research setting. • Stiffness covers time to movement and can be good or bad depending on previous day's activities and is measured over the previous week (covered in BASDAI for 0–2 hours). Evidence of a trend over time would be expected as an outcome and therefore a reported 30 minutes of difference is not necessarily clinically meaningful. • No effect of acupuncture on peripheral outcomes were reported in the studies • Schober and finger-to-floor distances outcomes do not vary together in the way that would be expected, which makes the results difficult to interpret.
Other considerations	The GDG noted that there is no standardisation for acupuncture in the UK so it is unclear if the methods of acupuncture described are generalisable to UK practice. The GDG agreed that based on the evidence presented that they were unable to recommend the use of acupuncture for the treatment of spondyloarthritis. The GDG agreed there is likely to be value in randomised controlled trials looking at the effectiveness of acupuncture. A definitive trial would either (if positive) enable an effective intervention to be introduced or (if negative) allow money to be saved by withdrawing ineffective services. It was agreed that any trials conducted should have a control arm of sham acupuncture, as these have a considerably lower risk of bias than trials using standard care as a comparator.

6.4.5 Recommendations

No recommendations were made.

6.4.6 Research recommendations

11. What is the effectiveness and cost-effectiveness of acupuncture, as standardly performed in the UK, versus sham acupuncture for the management of symptoms in axial and peripheral spondyloarthritis?

Why this is important

Until recently many people with spondyloarthritis have received acupuncture as a treatment for pain in spondyloarthritis, as these treatments have been available through many NHS services. However, there is currently a lack of evidence of efficacy of acupuncture in this population. Therefore, if maintenance of access to acupuncture is going to be justified for

1   people with spondyloarthritis, well-conducted long-term randomised controlled trials (in both
2   axial and peripheral spondyloarthritis) comparing acupuncture plus standard care to standard
3   care alone are necessary, given the sparse and low-quality evidence base that currently
4   exists.

6.5 Physical aids for spondyloarthritis

Review question 18

- What is the effectiveness of physical aids (for example, braces) compared with standard care for managing spondyloarthritis?

6.5.1 Evidence Review

The aim of this review was to assess the effectiveness of physical aids for the management of symptoms (axial and peripheral) of spondyloarthritis. The review focused on identifying studies that fulfilled the conditions specified in Table 31.

Table 31: PICO table physical aids for spondyloarthritis

Population	Patients with spondyloarthritis
Interventions	Physical aids including braces, physical supports, splints, frames, stick
Comparators	Standard care including usual care (treatment as usual), waiting list, delayed start of treatment, no treatment, placebo exercise
Outcomes	Pain, adverse events, joint mobility, physical function, quality of life, imaging, composite measures

For full details of the review protocol please see Appendix C.

Randomised controlled trials were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well.

A systematic search and a hand search of the reference lists of systematic reviews identified 173 references. The references were screened on their titles and abstracts and 0 references were ordered for full text.

6.5.1.1 Description of included studies

No studies met the inclusion criteria for this review. Since no relevant studies were found, expert testimony was sought from 1 co-opted committee member and 1 external expert, which was presented as evidence to the committee (see Appendix I for a summary of this evidence).

6.5.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

6.5.3 Evidence statements

No evidence was identified

6.5.4 Evidence to recommendations

Relative value of different outcomes	The committee prioritised outcomes of pain, mobility and function (including composite measures), as these were the outcomes felt most likely to be improved with these interventions and of value to people with spondyloarthritis. Other relevant outcomes specified included quality of life, finger-floor distance and morning stiffness.
Trade-off between	Physical aids were discussed by the committee as having a role in

benefits and harms	both symptom management and assisted daily living.. Occupational therapy assessments can encompass more dimensions than those for which the initial referral was made e.g. full assessment of person's daily needs, beyond symptom management. However the committee noted that access to occupational therapists is not available for all patients due to variations in availability between areas.
	The use of back braces in ankylosing spondylitis may be contraindicated, particularly during inflammation Lack of access to physical aids could limit participation in day to day activities, which could have a negative impact on both physical and mental wellbeing.
Economic considerations	No studies were found in the health economic review and no de novo modelling was conducted for this review question.
	The GDG noted that the provision of occupational therapy is an overarching NHS issue, and access to appropriate therapy is predominantly influenced by factors beyond the control of referring clinicians. The group agreed that, if services are judged to be worth commissioning generally, people with spondyloarthritis represent a good example of people who would derive substantial benefit from them, in the ways identified in the therapists' statements. For this reason, referrals can be assumed to provide reasonable value for money. However, the group acknowledged that it had no formal evidence as to the cost-effectiveness of these services for people with spondyloarthritis, and this uncertainty was reflected in the strength of the GDG's recommendation – that referral should be considered rather than universally offered.
Quality of evidence	No RCT evidence was identified.
	Written statements were received from two occupational therapists outlining current assessment processes and specific physical aids made available on an as-needed basis. The GDG agreed with and supported the evidence presented noting that it was consistent with their knowledge and experience.
Other considerations	The GDG discussed access to occupational therapists, noting that access is not universal or standardised. Sometimes referral follows a social care assessment e.g. if a person has been in need of home based care or support. The GDG noted that access to assessment and support for driving aids (such as additional mirrors to correct for limited neck rotation) may be provided by the DVLA rather than occupational therapy, and that people with spondyloarthritis may need to initiate contact with the DVLA themselves.
	The GDG noted that supported and advice on the use of physical aids may be provided by a range of sources including various allied health care professions. It was therefore considered appropriate to not specify that this advice should be provided by a single group. It was noted that people with spondyloarthritis may have varying needs with regard to assistance to daily activities and that specialist assessment would be the best way to determine what kind of physical aids would be useful. The GDG agreed that people's needs for physical aids may change over time and that future review after initial assessment would be useful in supporting this. However, in the absence of evidence and acknowledging that individuals' needs may vary, it was agreed that it would not be appropriate to specify an exact frequency of review.
	Existing consensus-based recommendations from the NICE guidelines for management of rheumatoid arthritis were discussed. In the absence of any evidence from the literature, the GDG developed a recommendation based on their experience and expertise.

1   **6.5.5   Recommendations**

2   1.1.1   Consider a referral to a specialist therapist (such as a physiotherapist,

3   occupational therapist, hand therapist or podiatrist) for people with

4   spondyloarthritis who have difficulties with any of their everyday activities.

5   The specialist therapist should:

6   • assess their needs

7   • provide advice about physical aids

8   • arrange periodic reviews to assess the person's changing needs.

1   **6.5.5   Recommendations**

2   1.1.1   Consider a referral to a specialist therapist (such as a physiotherapist,

7 Surgical Interventions

Spondyloarthritides are progressive long-term medical conditions that are primarily managed by patient self-management, physiotherapy and rheumatology care including specialist disease-controlling medication. However, in selected circumstances, particularly where significant pain or physical restriction is not responsive to physiotherapy or medical interventions, orthopaedic intervention can play a crucially important role.

Patients with active inflammatory arthritis can develop joint damage severe enough to warrant joint replacement. This is thought to occur more commonly in patients with a history of inadequately controlled joint inflammation and tends to occur much earlier than the damage seen in degenerative osteoarthritis. Outcomes from joint replacement surgery have the potential to be affected by factors related to patient's arthritis. These include the impact of medications used to control joint synovitis – such as corticosteroids standard DMARDs or biological DMARDs – on perioperative healing and infection. Rehabilitation of patients with ongoing active inflammation in their joints or significant joint damage and physical restriction may be more challenging postoperatively.

In patients with axial spondyloarthritis and significant flexion deformities, vertebral wedge osteotomies may be undertaken to improve those deformities, particularly when forward vision is severely impaired. This is usually only after careful consideration by highly specialised spinal surgeons. Surgical intervention (often urgent) is also appropriate in axial spondyloarthritis patients with spinal fusion and a subsequent post-traumatic pseudoarthrosis of the spine, which can cause significant long-term neurological damage if left surgically untreated.

7.1 Predictors of surgical outcome

Review Questions 34 and 35

- What factors predict clinical improvement after spinal surgery (including osteotomy and fusion) in people with axial inflammation?
- What factors predict clinical improvement after joint replacement surgery?

7.1.1 Evidence review

The aim of this review was to identify the prognostic factors that predict clinical improvement following i) spinal surgery in axial inflammation and ii) joint replacement.

The review focussed on identifying studies that fulfilled the conditions specified in Table 32.

Table 32: PICO table: Predictors for surgical outcome

Population	People (aged 16 years and over) with a confirmed diagnosis of spondyloarthritis who have undergone or are going to undergo surgery
Variables	• duration of disease • duration of delay in diagnosis • severity of disease • comorbidities • osteoporosis • site of surgery • indication for surgery • elective/non-elective • current treatment • fitness for surgery • pre-surgical functional status • type of centre delivering surgery • smoking • NSAID use
Endpoint	Good (or poor) surgical outcome
Outcome	Predictors assessed on diagnostic test accuracy measures

For full details of the review protocol please see Appendix C.

Prospective consecutive case series are the preferred study type. If none were available, other case series – such as retrospective case series or case series where it is not clear if cases are consecutively recruited – were considered.

7.1.1.1 Studies identified

A systematic search and a hand search of the reference lists of systematic reviews identified 5,996 references. The references were screened on their titles and abstracts and 26 studies were ordered.

Twenty three studies were excluded as they did not meet the eligibility criteria, such as the condition of interest, or primarily because the study did not report any relevant or usable data. A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F.

No studies met the inclusion criteria for predictors of clinical improvement after spinal surgery. A total of 3 case series was included for predictors of clinical improvement after joint replacement surgery.

1 An additional 23 studies were ordered during the re-run searches, of which 22 were excluded
2 and 1 case series was included for predictors of clinical improvement after joint replacement
3 surgery.

4 **7.1.1.2 Description of included studies**

5 Evidence tables for included studies can be found in Appendix E, with GRADE profiles
6 reported in Appendix G

7 The included observational studies all investigated hip arthroplasty in people with ankylosing
8 spondylitis, examining the following surgical outcomes:
9 • the need for surgical revision
10 • postoperative flexion
11 • the need for transfusion
12 • heterotopic ossification and
13 • poor healing of the surgical incision

14 The predictive factors examined in relation to the above surgical outcomes included:
15 • age
16 • sex
17 • weight
18 • duration of symptoms
19 • steroid use
20 • bleeding acetabular protrusion
21 • ankylosis
22 • preoperative C-reactive protein levels
23 • preoperative erythrocyte sedimentation rate
24 • time between surgeries
25 • heterotopic ossification
26 • type of anaesthesia and
27 • type of femoral head.

28 **7.1.1.3 Variations from protocol**

29 There was a lack of studies identified that reported the desired diagnostic test accuracy data,
30 or from which the reviewer could calculate these. Therefore, studies reporting the
31 relationship between prognostic factors and surgical outcomes using other measures (such
32 as odds or hazard ratios) were also considered.

33 **7.1.1.4 Minimal clinically important differences**

34 When considering diagnostic accuracy data for individual factors in isolation, it was agreed in
35 collaboration with the GDG that a positive likelihood ratio of 2 would constitute significant
36 diagnostic value. Therefore, when interpreting the diagnostic accuracy results for single
37 factors in isolation, something would only be considered to have diagnostic value if the result
38 was statistically significant at the 95% confidence level, and the point estimate of the positive
39 likelihood ratio was greater than 2 (or equivalently, value at ruling out the disease if the
40 negative likelihood ratio was less than 0.5). Since the majority of data identified was not in
41 the form of diagnostic accuracy data (see variations for protocol section above) it was not
42 possible to consistently apply this rule across the different outcome measures.

43

7.1.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

7.1.3 Evidence statements

Predictors of arthroplasty revision due to loosening of prosthetic components

Very low-quality evidence from a single study reported that increasing age, weight, female sex, steroid use and bleeding greater than the mean were not significantly associated with outcome of surgery.

Predictors of poor postoperative function (hip flexion)

Very low- to low-quality evidence from a single study found that female sex, acetabular profusion, heterotopic ossification and ankylosis have a sensitivity of 51.5% or less in predicting poor postoperative outcomes. Ankylosis had poor specificity while female sex and acetabular profusion had a specificity of 86.8% or more. Use of a 32 mm femoral head had sensitivity and specificity at around 75%.

The same study reported that, after multivariate regression, higher preoperative C-reactive protein levels and increased levels of heterotopic ossification were protective (OR<1) against poor hip flexion while the use of a 32 mm femoral head was predictive of poor hip flexion (OR>1)

Predictors of blood loss

Very low-quality evidence from a single study reported that being underweight (defined as BMI less than 18.5 kg/m^2) had a sensitivity of 43.3% and a specificity of 74.6% in predicting blood loss.

Predictors of heterotopic ossification

Very low-quality evidence from a single study found that female sex, preoperative hip ankylosis, heterotopic ossification in previous total hip arthroplasty, elevated preoperative C-reactive protein levels and elevated erythrocyte sedimentation rates increased the risk of heterotopic ossification.

7.1.4 Evidence to recommendations

Q34: Predictors of clinical improvement following spinal surgery

Relative value of different outcomes	The GDG identified that people with axial spondyloarthritis with possible indications for spinal surgical intervention fall into 3 categories:
	• Those with severe kyphosis (that is, people who have developed a 'question mark' posture).
	• Those with stable fractures e.g. caused by osteoporosis
	• Those with potentially unstable fractures caused by trauma
	Outcomes likely to be of most importance to people with axial spondyloarthritis with an indication for spinal surgery are:
	• Functional improvement
	• Improved quality of life – gaze, balance, gait

	• Prevention of further deformity (kyphosis, stable fractures)
	• Prevention of spinal/neurological damage (potentially unstable fractures).
Trade-off between benefits and harms	The GDG discussed 2 spinal surgical procedures: osteotomy and fusion to correct deformity/restore sagittal balance, and fusion to treat fractures.
	Regarding benefits, there is a distinction between different indications:
	• For severe deformity or stable fractures the benefits will relate primarily to quality of life issues e.g. functional impairment
	• For post-traumatic unstable fractures there is a risk of neurological damage if these remain untreated.
	Risks of osteotomy are high at present (the experience of the GDG suggested a 10% mortality rate, 30% risk of neurological complications of varying degrees of severity [ranging from nerve root injury to quadriplegia], additional risks of gastrointestinal complications or aortic injury).
	Risks arising from osteotomy may also be influenced by the operator/centre level of experience, as well as surgical decisions.
	The GDG noted that NHS England currently commissions 'complex spinal surgery services' through its specialised commissioning framework, and it agreed that services provided in this way were more likely to have the necessary expertise to carry out these procedures.
	Risks of fusion were not quantified but it was noted by the co-opted surgical expert that interventional spinal fusion is more complex in people with axial SpA due to ankylosis of the spine.
	Because the risks of spinal surgery are high in people with spondyloarthritis, and there are non-surgical options to prevent spinal deformity (e.g. physiotherapy and biological DMARDs), the GDG agreed it would only be appropriate to refer people for surgery once these options had been exhausted.
Economic considerations	No studies were found in the health economic review and no de novo modelling was conducted for this review question. The GDG did not believe that its recommendations would incur any new opportunity costs, as the circumstances under which referral is warranted are limited and reflective of current practice.
Quality of evidence	No evidence was found which fulfilled the requirements of the review protocol. Studies were excluded for a number of reasons including failing to provide adequate quantitative information to inform the pre-specified outcome measures agreed by the GDG.
	Two relevant expert members of the GDG (a spinal surgeon and a radiologist) both highlighted the risks and benefits of the surgery, and the difficulties of radiographic detection of fractures which may indicate need for surgery.
Other considerations	It was acknowledged that the scope of the question was to identify predictors of outcome from spinal surgery, though the GDG's discussions and consensus statement in the absence of evidence on this topic focused mostly on indications for referral and assessment.
	It was noted that surgical intervention to address deformity in people with axial spondyloarthritis is considered a high risk procedure and is indicated when optimal non-surgical therapy (including physiotherapy) has failed to halt progression.
	It was noted that, in the co-opted spinal expert's experience, surgical intervention at an earlier stage in people with axial-spondyloarthritis whose spinal deformity is likely to progress to a severe degree may have a lower risk of complications from osteotomy. The GDG noted that this should not be at the expense of pursuing optimal non-

	surgical therapy (including physiotherapy) as the primary treatment strategy.
	It was noted that people presenting with spinal fractures are likely to have additional fractures at other spinal locations which may not be detected during investigation/assessment of the index fracture.
	It was discussed that there is a need for rheumatology and surgical teams to work together in the pre- and post-surgical management of people with SpA undergoing spinal surgery.

1   **Q35: Predictors of clinical improvement following joint replacement surgery**

Relative value of different outcomes	The GDG discussed whether a dichotomous measure of 90° hip flexion was a useful measure. It was noted that dichotomous measures may reflect culturally specific standing/sitting practice, but that this cut-off was appropriate for a UK population.
	Post-operative heterotopic ossification was considered by the GDG to be an outcome of surgery rather than a predictor of good/poor surgical outcome, though there was not any available evidence to evaluate this.
Trade-off between benefits and harms	The GDG noted that the harms presented in some of the papers may be indicative of clinical practice that may differ from that conducted in the UK. For example, Zhao *et al.* reported a high number of people receiving blood transfusion during surgery. It was also considered that the post-operative surgical management presented in one paper was unclear and may differ to that in the UK.
Economic considerations	No economic evidence was presented
Quality of evidence	Two studies were retrospective and one was unclear on this point. It was unclear as to whether cases were recruited consecutively. The reference standard to which predictors were being compared was not reported. Population characteristics were possibly under-reported. The pre-specified outcome measures (specificity, sensitivity etc.) were not reported in the studies and had to be calculated from available data by the analyst.
	The studies reported data on the following outcomes:
	• Predictors of arthroplasty revision due to loosening of prosthetic components (Lehtimaki) (predictors: age, female sex, weight, steroids, bleeding>median)
	• Poor post-operative function (flexion) (Zhang) (predictors: female sex, acetabular protrusion, ankylosis, heterotopic ossification, use of a 32 mm femoral head, pre-operative C-reactive protein level)
	• Blood loss (need for transfusion) (Zhao) (predictors: underweight (BMI<18.5))
	• Predictors of poor healing of surgical incision (Zhao) (predictors: underweight (BMI<18.5))
	CRP was discussed as a predictor. Although it is not an indicator for referral for surgery, the evidence presented may reflect lower CRP values occurring in people in whom disease is better controlled. It was also noted that (i) CRP may be elevated in a range of conditions e.g. infections and (ii) not everybody with active spondyloarthritis will have elevated CRP. It was noted that prior to surgery, DMARDs may have been stopped, and therefore the pre-operative CRP level may not reflect the typical level for a given person.
	In one study (Zhang), the patients included had a lower age than may be typical for referral for surgery in the UK. Younger people may be expected to have better outcomes from joint surgery compared with older people. Conversely, people with indication for joint replacement at a young age may have been experiencing severe disease with onset in childhood/adolescence. The data reporting in this paper was inconsistent and of poor quality, making it difficult to draw any

	meaningful conclusions. One study (Lehtimaki) – population: lots treated with steroids (>40% compared with an expected 10%), and had a high prevalence of amyloidosis (9.3%)]. The GDG considered this to be notably different to the typical UK spondyloarthritis population. The GDG agreed that there was insufficient evidence presented in order to make a clinical recommendation.
Other considerations	The GDG noted that most of the studies were conducted in Asian countries where the study population may differ sufficiently from the UK SpA population to make it difficult to extrapolate some of the findings from the presented evidence, and hence the GDG did not feel this evidence was sufficiently robust to be able to make recommendations based upon it. It was noted that some of the predictors reported in the included studies were actually peri-/post-operative factors which therefore could not be considered to be prognostic. It was discussed that there is a need for rheumatology and surgical teams to work together in the pre- and post-surgical management of people with SpA undergoing joint replacement surgery.

7.1.5 Recommendations

20. Do not refer people with axial spondyloarthritis to a complex spinal surgery service to be assessed for spinal deformity correction unless the spinal deformity is:

- **significantly affecting their quality of life and**

21. severe or progressing despite optimal non-surgical management (including physiotherapy).If a person with axial spondyloarthritis presents with a suspected spinal fracture, refer them to a specialist to confirm the spinal fracture and carry out a stability assessment. After the stability assessment, the specialist should refer people with a potentially unstable spinal fracture to a spinal surgeon.

7.1.6 Research recommendations

12. Is pre-operative disease activity/stability a predictor of outcomes after spinal surgery for people with spondyloarthritis and axial inflammation?

Why this is important

Spinal surgery is only considered for a small subset of people with spondyloarthritis. To maximise the benefit-risk balance from surgery, it is necessary to identify in advance those individuals who will gain the greatest benefit, which in turns requires evidence linking pre-surgical characteristics to outcomes. Pre-operative disease activity is felt to be one of the factors most likely to correlate to surgical outcomes, but there is currently no evidence to support or refute this belief. Cohort studies (either prospective or retrospective) would provide evidence which could help to identify cut-offs for the appropriate people to refer for surgery, which is not currently possible.

13. Is pre-operative disease activity/stability a predictor of outcomes after joint replacement surgery for people with spondyloarthritis?

1 **Why this is important**

2 Joint replacement surgery is only considered for a small subset of people with
3 spondyloarthritis. To maximise the benefit-risk balance from surgery, it is necessary to
4 identify in advance those individuals who will gain the greatest benefit, which in turns
5 requires evidence linking pre-surgical characteristics to outcomes. Pre-operative disease
6 activity is felt to be one of the factors most likely to correlate to surgical outcomes, but there
7 is currently no evidence to support or refute this belief. Cohort studies (either prospective or
8 retrospective) would provide evidence which could help to identify cut-offs for the appropriate
9 people to refer for surgery, which is not currently possible.

8 Organisation of care and long-term monitoring

Once a diagnosis of spondyloarthritis has been made, a treatment plan can be developed, comprising pharmacological interventions, access to physiotherapy services, and advice on exercise and self-care. The spondyloarthritides are chronic conditions and therefore need ongoing management. This may involve monitoring the effectiveness of different interventions, and supporting people to manage their condition as their life needs change.

Different regions may organise care provision differently, which may present either benefits or challenges to people seeking support. In some places, the ongoing management of spondyloarthritis may be performed very effectively in primary care, but in other areas where appropriate knowledge and services are lacking, it would be preferable for people to receive all their ongoing care and support in specialist settings.

It is not uncommon for people with spondyloarthritis to have other chronic co-morbid conditions requiring long-term management. The cross-speciality management of these conditions is important, particularly where treatments overlap. For example, people with inflammatory bowel disease as well as spondyloarthritis may need the gastroenterology team in charge of their care to liaise with rheumatology services in order to optimise a biological DMARD dose that manages both conditions effectively. Systems that support cross-disciplinary communication should therefore be in place.

People with spondyloarthritis may experience 'flares' during which their symptoms are exacerbated, which may warrant additional management with increased or additional doses of pharmacological treatments, or access to physiotherapy services. Different models of care during flare episodes may suit different people, with some being able to self-manage with advice, some needing assistance in primary care, while others may need access to a rheumatology team. Organising health service access in such a way to enable needs and preferences of the person experiencing a flare episode to be considered is therefore important.

The treatments used in spondyloarthritis may have side effects or lead to the development of complications over time. Pharmacological treatments in particular may be associated with complications in many populations, and were therefore evaluated in the context of spondyloarthritis. Being a group of systemic inflammatory conditions, the spondyloarthritides themselves may be associated with the development of co-morbidities or complications over time. Both disease manifestations and responses to treatment therefore need to be monitored throughout the disease course.

8.1 Transition from paediatric to adult rheumatology services

Review Question 13

- How should transition from specialist paediatric services to specialist adult rheumatology services be managed for young people between the ages of 16 and 18?

8.1.1 Evidence review

The aim of this review was to describe how care for young people with spondyloarthritis should be managed as they transition to adult services.

Following the publication of NICE Guidance NG43 'Transition from children's to adults' services for young people using health or social care services', the GDG opted not to request an additional evidence review specific to spondyloarthritis for this topic, as it was not felt that there would be any further condition-specific issues to consider.

8.1.2 Recommendations

22. For guidance on managing the transition of young people with juvenile idiopathic arthritis to adult services, see the NICE guideline on transition from children's to adults' services for young people using health or social care services [NG43]

8.2 Monitoring of pharmacological interventions used in spondyloarthritis

Review Question 22

- How often should people receiving pharmacological interventions for managing spondyloarthritis be monitored? How often should people with spondyloarthritis be offered specialist review?

8.2.1 Evidence review

The aim of this review was to assess the frequency of monitoring for those receiving pharmacological interventions and frequency of specialist review for all of those with spondyloarthritis.

The review focussed on identifying studies that fulfilled the conditions specified in Table 33.

Table 33: PICO table monitoring

Population	People (aged 16 or over) with a confirmed diagnosis of spondyloarthritis
Interventions	Frequency of medication monitoring or review
Comparators	No monitoring, different monitoring strategies
Outcomes	Outcomes for frequency of monitoring tolerability, adverse events, adherence Outcomes for specialist review monitoring standard outcomes for SpA intervention reviews

For full details of the review protocol please see Appendix C.

The study designs specified in the protocol for addressing this review question were systematic reviews and randomised controlled trials. A total of 5,944 references were identified for this question, of which 8 were ordered for full-text review. All of these studies were excluded from the final review, with the reasons for exclusion given in Appendix F.

8.2.1.1 Description of included studies

No studies met the inclusion criteria for this review.

8.2.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question

8.2.3 Evidence statements

No evidence was identified

8.2.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that it was important to monitor people with spondyloarthritis who are receiving pharmacological therapies so as to be able to check ongoing efficacy or note the emergence of adverse events or complications. They discussed that the frequency of monitoring varied depending on which classes of drugs were being taken.
Trade-off between benefits and harms	The GDG noted the general risks of long term use of non-steroidal anti-inflammatory drugs (e.g. cardiovascular events, gastrointestinal

	bleeding) and the need to consider protein-pump inhibitor co-prescription and monitoring of renal function using eGFR measurements. he GDG discussed issues relating to shared care agreements and noted that there can be issues in how these are managed in practice. It noted that some regions do not establish such agreements. Where they do exist, issues include requested blood tests not being carried out, and people receiving live vaccinations, or other contra-indicated interventions, while receiving biological DMARDs. Communication between primary and secondary care was deemed to be important.
Economic considerations	No health economic evidence was found and this review question was not prioritised for health economic modelling. The recommendations made were not thought to be associated with any new opportunity costs.
Quality of evidence	No studies met the eligibility criteria for the review. The GDG discussed issues relating to the review question informed by their expertise.
Other considerations	The GDG agreed to cross-refer to other generic sources of NICE Guidance such as the Medicines Optimisation guideline. Other sources of information that the GDG considered to be reliable were the British Society of Rheumatology's guidance on monitoring of disease-modifying anti-rheumatic drug (DMARD) use, as well as individual Specific Product Characteristics which advise on how often and what type of tests should be used to monitor drug response. The GDG felt that it was appropriate to draw attention to the need to take into account the general adverse event profiles of NSAIDs and DMARDs when prescribing them in this population. The GDG opted to make a recommendation in the absence of direct evidence, based on their clinical experience and expertise. It was agreed that the recommendation should be presented alongside other recommendations on the organisation of care for people with spondyloarthritis (see section 8.5.5)

8.2.5 Recommendations

23. **For guidance on monitoring long-term pharmacological treatments, see the NICE guideline on medicines optimisation.**

24. **Take into account the adverse effects associated with NSAIDs, standard DMARDs and biological DMARDs when monitoring spondyloarthritis in primary care.**

8.2.6 Research recommendations

14. **What are the most effective doses and monitoring arrangements for people treated with anti-tumour necrosis factor (TNF) drugs both for spondyloarthritis as well as a comorbidity (e.g. inflammatory bowel disease) simultaneously?'**

Why this is important

Anti-TNF therapy is indicated for a number of conditions, and it is therefore not uncommon for people to be treated with anti-TNFs for more than one condition simultaneously. This means it is not possible to follow the optimum dosing or monitoring strategy for both conditions, as these will frequently be different, leading to uncertainties in the correct management for that individual. There is therefore the need for studies of people on anti-TNF therapy for spondyloarthritis and a common anti-TNF treated comorbidity (e.g. inflammatory

1   bowel disease) to identify the optimum treatment arrangements for each relevant pair of
2   conditions. Trials would need to measure outcomes relevant to both conditions as well as
3   overall health-related quality of life to measure the trade-offs between optimal control of each
4   condition.
5

8.3 Care setting for management of flare episodes

Review Question 29

- What is the usefulness of direct access to specialist care, compared with initial primary care access followed by specialist rheumatological care, in the management of flare episodes?

8.3.1 Evidence Review

The aim of this review was to assess the usefulness of direct access to specialist care in the management of flare episodes.

The review focussed on identifying studies that fulfilled the conditions specified in Table 34.

Table 34: PICO table specialist care for flare episodes

Population	Patients with diagnosed spondyloarthritis who have flare episodes
Interventions	Care by health care professional in specialist setting
Comparators	Care by health care professional in primary care settings followed by specialist care
Outcomes	Time to care received, number of contacts with health care professionals, satisfaction with care received, health-related quality of life, resource use and cost, improvement in severity, duration, frequency of flare episodes

For full details of the review protocol please see Appendix C.

Potential study designs that may have helped to answer this question and were in the review protocol included RCT, observational intervention and qualitative study designs.

A systematic search identified 956 references. The references were screened on their titles and abstracts and 2 references were ordered for full text. Both of these studies were subsequently excluded (see Appendix F for reasons).

8.3.1.1 Description of included studies

No studies met the inclusion criteria for this review

8.3.2 Health economics evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question

8.3.3 Evidence statements

No evidence was identified

8.3.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that it was important to consider direct access to specialist care, compared with initial primary care access followed by specialist care, in the management of flare episodes. The GDG discussed the provision of care for the management of flare within primary care. It was noted that there is currently a wide variety in the level of service provision, with some primary care services providing management themselves while others referring

	patients directly to specialist care. Due to the current variations in practice the GDG discussed the need to ensure that there is a balance between providing beneficial management in primary care where there is sufficient expertise and ensuring that where it is needed review by specialist services is undertaken.
Trade-off between benefits and harms	The GDG noted that the difficulties with defining flare make providing definitive recommendations in this area difficult. Individual patients will experience flare episodes in different ways with a variety of symptoms making the provision of a generic management strategy challenging. It was noted that other inflammatory triggers (for example, infection) may resemble flare and would also necessitate assessment.
Economic considerations	No health economic evidence was found and this review question was not prioritised for health economic modelling. The absence of economic evidence was one of the considerations that precluded the GDG from making prescriptive recommendations in this area. The group concluded that local health economies would be best-placed to establish efficient arrangements for their patients, and that the GDG could only specify the critical principles these should cover.
Quality of evidence	There was no evidence identified from the searches that considered the usefulness of direct access to specialist care, compared with initial primary care access followed by specialist care, in the management of flare episodes. The study designs considered eligible for this question included observational studies and qualitative studies.
Other considerations	The GDG noted the importance of recognising flare and preventing any further possible disease progression or complication. The GDG further noted the need to discuss with people with spondyloarthritis their flare episodes particularly where this may help identify other contributing factors that may trigger the onset of flare episodes. The GDG agreed that where individualised care plans for flare management can be agreed between care services then management of flare within primary care services can be effective. However it was agreed that persistent or recurrent flare episodes should be reviewed by specialist services. The GDG discussed the management of flare in primary care services. They considered that primary care was appropriate for some episodes of care but that there are specific cases where specialist services should be used. These included uveitis, patients taking biological DMARDs, those with spondyloarthritis-associated comorbidities, those needing multidisciplinary healthcare support, and cases of recurrent or persistent flare. The GDG noted that uveitis may occur during a flare episode, and this should result in same-day review by an ophthalmologist. The GDG considered that using their expertise and clinical experience they could develop consensus-based recommendations for this area.

1   **8.3.5   Recommendations**

2   **25. Manage flares in either specialist care or primary care depending on the person's**
3   **needs.**

4   **26. When managing flares in primary care, seek advice from specialist care as**
5   **needed, particularly for people who:**

6   • have recurrent or persistent flares

- are taking biological DMARDs
- have comorbidities that may affect treatment or management of flares.

27. Be aware that uveitis can occur during flare episodes. See recommendation 5.1 for guidance on immediate (same-day) ophthalmological assessment for people with acute anterior uveitis.

8.3.6   Research recommendations

15. What is the comparative effectiveness and cost-effectiveness of direct access to specialist care versus access via primary care for reducing the risk of complications during flare episodes?

Why this is important

There is currently no evidence about the optimal setting for managing flares and the most appropriate route for accessing specialist care, and there is considerable variation in practice across the UK. Cluster randomised RCTs comparing direct access to specialist care versus access to primary care could enable a greater standardisation of services, by demonstrating which of these two outcomes produces better outcomes for individuals experiencing a flare.

16. What is the comparative effectiveness and cost-effectiveness of healthcare professional led management and self-help plans for the management of flare episodes in people with spondyloarthritis?

Why this is important

There is currently no evidence about the relative effectiveness of self-management versus healthcare professional management of flares in people with spondyloarthritis, and there is considerable variation in practice across the UK. RCTs comparing these two approaches (which would need to follow people up for at least the entire duration of their flare) could help to demonstrate whether there are additional benefits of healthcare professional led management which would justify the higher costs of such an approach.

8.4 Care setting for long-term management

Review Question 30

- What is the effectiveness of specialist-led long-term management of spondyloarthritis compared with primary-care-led long-term management?

8.4.1 Evidence review

The aim of this review was to assess the care setting for long-term management of spondyloarthritis.

The review focussed on identifying studies that fulfilled the conditions specified in Table 35.

Table 35: PICO table care setting for long term management

Population	Patients with diagnosed spondyloarthritis
Interventions	Specialist-led management
Comparators	Primary-care-led management
Outcomes	Number of contacts with health care professionals
	Number, severity, duration of flare episodes
	Resource use and costs
	Health-related quality of life
	Disease progression
	Long term morbidity and extra-articular symptoms and mortality (including but not limited to uveitis, psoriasis, inflammatory bowel disease, enthesitis, oligoarthritis, site specific inflammation, dactylitis, osteoporosis (and fracture), spinal fractures, spinal cord injuries, blindness, aortic regurgitation, cardiovascular complications, joint replacement)
	Access to different therapy options (including, but not limited to, drug therapies)
	Access to specialist therapies (e.g. specialist rheumatology physiotherapy)

For full details of the review protocol please see Appendix C.

Potential study designs that may have helped to answer this question and were in the review protocol included RCT and observational intervention.

A systematic search identified 5,304 references. The references were screened on their titles and abstracts and 2 references were ordered for full text. Neither of the ordered studies met the review protocol inclusion criteria the excluded studies list is available in Appendix F.

8.4.1.1 Description of included studies

No studies met the inclusion criteria for this review

8.4.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

8.4.3 Evidence statements

No evidence was identified

1 **8.4.4 Evidence to recommendations**

Relative value of different outcomes	The GDG agreed the importance of access, for people with spondyloarthritis, to specialist and more general care during the long-term management of their condition. It was agreed that where this access may be fragmented that this could impact negatively on the quality of life, flare episodes and other aspects of the long-term management of their condition for those with spondyloarthritis.
Trade-off between benefits and harms	The GDG noted that no evidence had been identified for this review question, therefore no conclusions could be drawn from evidence on whether specialist-led or primary care-led long-term management is more effective.
	The GDG noted that the division between the settings within this review question may be somewhat arbitrary. The GDG discussed whether there may be some variability or inconsistency in the nature of primary care-led management depending on the specialism available in the general practice. This can mean that the setting for long-term care may be less significant than the availability of specialist knowledge relating to spondyloarthritis. The need for specialist secondary care input for some aspects of care, such as the prescribing of certain treatments was noted. The GDG agreed that it is important for people with spondyloarthritis to be aware of which service to contact if there are changes in their condition, that long-term management throughout the lifetime of a chronic condition is likely to require a combination of care across different settings.
	In consideration of the absence of evidence and reflecting on their discussion of this review question, the GDG made a consensus based recommendation.
Economic considerations	No health economic evidence was found and this review question was not prioritised for health economic modelling.
	As in sections 8.3 and 8.5, the GDG concluded that local health economies would be best-placed to establish efficient arrangements for their patients and service, so it did not make a prescriptive recommendation.
Quality of evidence	No evidence was identified via the search undertaken for this question.
Other considerations	The GDG opted to make a consensus-based recommendation based on their experience.

2 **8.4.5 Recommendations**

3 **28. Ensure that people with spondyloarthritis have access to specialist care in**
4 **primary or secondary settings throughout the disease course to ensure optimal**
5 **long-term spondyloarthritis management.**
6

8.5 Cross-speciality care

Review Question 31

- How should cross-speciality care for people with spondyloarthritis be organised?

8.5.1 Evidence review

The aim of this review was to assess the organisation of cross-speciality care. The review focussed on identifying studies that fulfilled the conditions specified in Table 36.

Table 36: PICO table cross-speciality care

Population	Patients with diagnosed spondyloarthritis
Interventions	Cross-speciality care, which could include: • Combined clinics • Cross-speciality referrals • Cross-speciality treatment management • Multiple drug management
Comparators	n/a
Outcomes	Time to appointment, number of contacts with health care professionals, health related quality of life, resource use and costs, patient satisfaction, disease burden reduced from both spondyloarthritis and associated conditions, service delivery/organisation

For full details of the review protocol please see Appendix C.

Potential study designs that may have helped to answer this question and were in the review protocol included observational intervention study designs.

A systematic search identified 5,304 references. The references were screened on their titles and abstracts and 2 references were ordered for full text. Neither of the ordered studies met the review protocol inclusion criteria, with reasons for the exclusion given in Appendix F

8.5.1.1 Description of included studies

No studies met the inclusion criteria for this review

8.5.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

8.5.3 Evidence statements

No evidence was identified

8.5.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that the way that treatment and monitoring is organised across specialities can have a significant influence on outcomes and quality of life for people with spondyloarthritis. The appropriate use of health service resources was also considered to be important.
Trade-off between benefits and harms	The GDG noted that no evidence had been identified for this review question therefore no conclusions could be drawn from evidence on

	the organisation of cross-speciality care.
The GDG further noted that there was no evidence identified that related to the use of multidisciplinary teams in the care of those with spondyloarthritis.
The GDG agreed that central to any discussion about the organisation of care is the need for clear and effective communication across all care areas, both specialist and generalist. The fluctuating nature of spondyloarthritis conditions with comorbidities and extra-articular manifestations may necessitate changing or combining specialty input at differing stages of the condition. The GDG discussed the use of cross-specialty measures such as shared care plans or other forms of shared documentation and joint speciality clinics. The GDG noted that local service arrangement and commissioning often differ throughout the UK. It was agreed that whatever the local arrangements the most important aspect of cross-specialty care is the facilitation of communication between the care settings that are involved. The GDG discussed concerns that currently there is insufficient communication between specialist secondary care areas and that this can affect the co-ordination of care.
The GDG discussed the vital importance of providing ongoing access to multidisciplinary expertise for those with spondyloarthritis. They also noted the potential for harm or sub-optimal care if people who have spondyloarthritis and other co-morbidities do not have their care properly co-ordinated across specialities.
Economic considerations	No health economic evidence was found and this review question was not prioritised for health economic modelling.
Although the GDG felt strongly that well defined arrangements for coordinating care should be in place, it had no evidence as to a single optimal structure, so the recommendations it made were intended to allow local health economies to define suitable arrangements for their population and service.
Quality of evidence	No evidence was identified via the search undertaken for this question.
Other considerations	The GDG opted to make consensus based recommendations according to their experience. They noted that for co-ordination across primary and secondary care systems to take place, there was a need for facilitation from commissioners; without this, they felt it would be difficult for different specialities to put in place the necessary arrangements. The GDG opted not to specify exactly how things should be arranged at the local level, acknowledging that local needs and availability of resources would influence this. They nonetheless felt that communication between specialists and departments was an important element of co-ordination of care.
In light of the potential harms or suboptimal care people with spondyloarthritis with additional comorbidities may experience, the GDG considered it appropriate that a consensus-based recommendation be made to cover this situation. They noted that this was particularly relevant where people were taking standard or biological DMARDs for more than one condition.

8.5.5 Recommendations

29. Commissioners should ensure that local arrangements in place to coordinate care for people across primary and secondary care:

- prescribing NSAIDs and standard DMARDs
- monitoring NSAIDs, standard DMARDs and biological DMARDs

- managing flares
- ensuring prompt access to specialist rheumatology care when needed
- ensuring prompt access to other specialist services to manage comorbidities and extra-articular symptoms.

30. Ensure that there is effective communication and coordination between all healthcare professionals involved in the person's care, particularly if the person has comorbidities or extra-articular symptoms.

31. Ensure that there is communication and coordination between rheumatology and other relevant specialities (such as dermatology, gastroenterology and ophthalmology). This is particularly important for people who:

- are already receiving standard DMARDs or biological DMARDs for another condition
- need to start taking standard DMARDs or biological DMARDs for another condition.

8.6 Complications of spondyloarthritis

Review Question 32

- What are the complications associated with spondyloarthritis?

8.6.1 Evidence review

Patients with spondyloarthritis may be at risk of a number of extra-articular and long-term complications.

The aim of this review was to identify the complications associated with spondyloarthritis so that these can be added to information given to patients and can be monitored for in any regular patient review, managing the risk where appropriate.

Table 37: PICO table – complications of spondyloarthritis

Population	People (aged 16 years and over) with a confirmed diagnosis of spondyloarthritis
Complications	Osteoporosis, uveitis (anterior), Inflammation of the aorta/aortic valve, aortic regurgitation, psoriasis, inflammatory bowel disease, spinal fractures, spinal cord injuries, cauda equina syndrome, erectile dysfunction, restrictive pulmonary disease, ischemic heart disease, stroke/CVA, joint replacement, hyperlipidaemia/metabolic syndrome, surgery, major depression, alcoholism Hospitalisation for the above or for disease symptoms, spinal/joint deformity
Comparators	People who do not develop the above complications
Outcomes	Rates of each complication at pre-defined time points

A single systematic search was conducted for both question 32 (complications of condition) and question 33 (complications of treatments), which identified a total of 13,303 references. Cohort studies were considered to be the optimal study design for these questions. The references were screened on their titles and abstracts and 179 studies were ordered for full text for question 32 of which 165 were available at the time of submission for consultation. 160 studies were excluded as they did not meet the eligibility criteria such as complications not specified in the scope, or studies without a defined follow up period. A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F. A total of 5 studies were included in the original review. Following a modification of the protocol (see section 0) and an updated literature search at the end of the process, a further 13 studies were added to this review.

8.6.1.1 Description of included studies

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G

Ischemic heart disease

Five studies reported data on ischemic heart disease. Two studies (Brophy et al., 2012, Edson-Heredia et al., 2015) retrospectively measured acute myocardial infarction incidence in people with ankylosing spondylitis and psoriatic arthritis respectively. One retrospective study (Chou et al., 2014) looked at acute coronary syndrome in ankylosing spondylitis. One study (Haroon et al., 2015) retrospectively measured vascular death rates in ankylosing spondylitis. The final study (Hung et al, 2016) prospectively looked at coronary heart disease incidence in ankylosing spondylitis.

Aortic valve insufficiency

Two studies derived from the same cohort reported on aortic valve insufficiency. One study (Jantti 2002), reported 2 cases (10.5%) detected at 23 year follow up in people with psoriatic

arthritis. The other (Kaarela 2009), reported 1 patient who required aortic surgery and died of aortic valve insufficiency during follow up, though it is not clear if this was during the 8 year or 20 year follow up period.

Stroke/cerebrovascular events

Five studies reported data on cerebrovascular outcomes. One study (Zoller et al., 2012) reported rates of ischemic and haemorrhagic stroke in people with reactive arthritis in a prospective study. One study (Edson-Heredia et al., 2015) retrospectively measured incidence of stroke in people with psoriatic arthritis. Four studies (Brophy et al.,2012, Hung et al., 2016, Keller et al., 2014, Zoller et al., 2012) looked at events in people with ankylosing spondylitis: Brophy 2012 retrospectively measured incidence of cardiovascular disease and stroke as a composite measure, Hung prospectively assessed cerebrovascular disease, Keller retrospectively measured stroke incidence, and Zoller prospectively examined incidence of both ischemic and haemorrhagic stroke, reporting rates separately for each.

Anterior uveitis (Iritis)

Three studies reported on occurrence of iritis. Two studies (Hart et al.,1986, Kaarela et al., 2009) reported cases in reactive arthritis, and two studies (Egberg et al., 2015, Kaarela et al., 2009) in ankylosing spondylitis. All data were prospectively collected except in Egberg 2015.

Fracture

Four studies reported on fractures in people with ankylosing spondylitis. One study (Maillefert, 2001) examined bone density in over a period of 24 months and reported fractures detected during that period. The second study (Kang, 2014) prospectively followed a cohort of 298 people with ankylosing spondylitis, of whom 287 were examined at 2 years and 131 at 4 years. The remaining studies (Weinstein et al., 1982, Munoz-Ortego et al., 2014) retrospectively assessed participants, each over a six year period, recording incidence of acute spinal fractures and clinical vertebral fractures respectively.

Osteoporosis

One study (Maillefert, 2001) examined bone density in ankylosing spondylitis over a period of 24 months. There were 54 participants with a mean disease duration at baseline of 12.4 years (SD8.6).

Inflammatory bowel disease

Two studies reported on cases of inflammatory bowel disease. One study (Edson-Heredia et al., 2015) retrospectively measured incidence of Crohn's disease in people with psoriatic arthritis, with a mean follow up of 3 years (SD=1.3). The other study (Mielants et al., 1995) prospectively measured occurrence of inflammatory bowel disease in people with spondyloarthritis, with a mean follow up of 5.7 years.

Depression

Two studies retrospectively collected data on depression. One study (Shen et al., 2016) looked at occurrence in people with ankylosing spondylitis, over a median period of 5.99 years. The other study (Edson-Heredia et al., 2015) assessed incidence in people with psoriatic arthritis, with a mean follow up of 3 years (SD=1.3)

Psoriasis

Two studies reported occurrence of psoriasis at multiple time points. One study (Jantti 2002) followed a group of people with seronegative oligoarthritis, measuring rates of psoriasis at 8

1 and 23 years. Participants had no more than 6 months of arthritis disease duration at
2 recruitment.

3 Another study (Theander, 2014) followed 197 people with psoriatic arthritis over a period of 5
4 years. Mean duration of disease at baseline was 11.6 months in men and 10.6 months in
5 women.

6 **Surgery (joint/tendon/spinal)**

7 One study (Kaarela 2009) reported surgery outcomes at follow up in reactive arthritis and
8 ankylosing spondylitis. Both groups were derived from a group of people with seronegative
9 oligoarthritis for up to 32 years, who had no more than 6 months disease duration at
10 recruitment. Variations from protocol

11 After initial presentation of studies limited to those with pre-defined time points at follow up,
12 the GDG were concerned about the lack of evidence included for many of the pre-specific
13 complications. They therefore requested that studies with follow up which was not of pre-
14 defined duration (e.g. person-years, survival analysis) also be included, if they met the other
15 eligibility criteria.

16 **8.6.1.2 Minimal clinically important differences**

17 Since the majority of studies identified in this review were non-comparative, no minimal
18 clinically important differences were considered.

19 **8.6.2 Health economic evidence**

20 A literature search was conducted jointly for all review questions in this guideline by applying
21 standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A
22 total of 9,970 references was retrieved, of which none were retained for this review question.
23 Health economic modelling was not prioritised for this review question.

24 **8.6.3 Evidence statements**

25 Very low-quality evidence from 18 studies reported rates of adverse events (cardiovascular,
26 uveitis, osteoporosis, inflammatory bowel disease, psoriasis, depression and fracture) for
27 people with spondyloarthritis. Risks of cardiovascular events, osteoporosis and fractures
28 were identified as being both higher in this population than the UK general population, and
29 potentially modifiable with appropriate advice and monitoring.

30 **8.6.4 Evidence to recommendations**

Relative value of different outcomes	The group agreed that, although all complications were of importance, cardiovascular risk, osteoporosis and fracture were the most important outcomes to monitor due to their likely frequency in this group and the severity of both the complications and failing to prevent or manage them in a timely fashion.
	Fractures are a significant complication. Acute back pain may indicate a fracture rather than a flare. This can lead to bone deformity and associated problems.
	There is a high risk of osteoporosis due to steroid use in this group of patients. Osteoporosis is also part of inflammatory process of SpA.
Trade-off between benefits and harms	The need for any X-ray should be balanced against the potential harms of radiation exposure. Spinal fractures can lead to bone deformity and may also be associated with increased morbidity; therefore, patients should report any acute back pain as soon as

	possible and have further investigations, such as X-rays.
Economic considerations	No studies were found in the health economic review and no de novo modelling was conducted for this review question. The GDG did not believe that its recommendations would incur any new opportunity costs, as they are broadly reflective of current practice.
Quality of evidence	Evidence was identified on a limited number of the pre-specified complications, and where data were available the committee agreed that the evidence in some cases was poor quality or lacking in detail .
	Psoriasis Two studies provided very low quality evidence, with poor quality of reporting in one study. The GDG noted that the evidence in one study was derived from people with psoriatic arthritis and therefore (i) would not necessarily be suitable for extrapolating to other spondyloarthritis populations as the rates were likely to be different and (ii) not necessarily helpful for this question as psoriasis can be classified as a disease manifestation of psoriatic arthritis, rather than a complication.
	Vertebral fracture There were 2 studies providing very low quality evidence; 1 primarily looked at bone density changes over 2 months with incidental report of fractures report, which was not the primary outcome of the study). The second study (Kang et al., 2014)) was well reported, but was graded as very low quality evidence, especially with regards to loss to follow up. It was also noted that in this study the investigators only undertook lumbar and pelvis radiographs, so any thoracic fracture may be missed. Patients often report acute pain, which they may interpret as a "flare", rather than being aware that the pain could be associated with a fracture, so there may potentially be many missed fractures. The studies assessed here are therefore likely to underestimate the actual incidence of fracture. The GDG noted that fractures can contribute to spinal deformity and are associated with higher morbidity in people with spondyloarthritis.
	Osteoporosis/Osteopenia One study of 54 people with ankylosing spondylitis was included. As the study did not report rates of osteoporosis in the cohort at baseline, but did record that there was no significant change at follow up, the GDG considered that this was more of a measure of prevalence than of incident cases. There was also a high drop-out rate in this study.
	Bone density measurement of spine was used to identify osteoporosis, which the GDG viewed to be an inadequate method of way of assessing osteoporosis. The method used to measure bone density in the lumbar spine is important. Osteoporosis was not a primary outcome for this study. Anterior-posterior dual-energy X-ray absorptiometry (DEXA) scans were interpreted and compared to a pre-defined scale; it was agreed that these are not a particularly reliable measure and that the gold standard measure of bone density is CT scanning with quantification of bone volume within vertebrae. The GDG highlighted that the European League Against Rheumatism (EULAR) recommend using lateral lumbar spine DEXA, but evidence for this is also very poor and there are radiation exposure issues. The GDG discussed how the method of measurement used may underestimate the number of people with osteopenia. The average age of people included in the study was 37.3 years, and it was noted that using a T score to identify osteopenia in a young population may

be inappropriate; they agreed that the bone density score make age-adjusted comparisons.

Surgery

One study (reported in two papers) reported surgical outcomes, in a population of people with reactive arthritis and ankylosing spondylitis. The people in this study had their spondyloarthropathy classified after recruitment; the investigators used loose criteria to recruit and formed criteria for sub-cohorts at later date. The GDG agreed that the prevalence of surgery in this population appeared to be in line with what they would have expected based on their clinical experience.

Aortic valve insufficiency

This outcome was reported in 2 papers and was rated as very low quality; one with a population of people with PsA and one with a mixed population of people with reactive arthritis and ankylosing spondylitis. It was noted that aortic valve insufficiency was not the primary outcome in either study. For one study with multiple follow up points, it was unclear at whether the incident cases were recorded at 8 or 20 years follow up. The GDG noted that the outcome was incidental in these papers and that aortic valve insufficiency is clinically silent, therefore the study may have underestimated the true prevalence.

Iritis/Uveitis

The GDG noted that iritis may be better classified as a co-morbidity of spondyloarthritis than as a complication. One very low quality study with a population of people with ankylosing spondylitis reported this outcome, which was assessed as very low quality. The study reported that there were 4 patients with multiple episodes of iritis. As with aortic insufficiency, it was unclear whether these incidents occurred at 8 or 20 years follow up.

Other considerations	The GDG noted that the outcomes specified originally in the protocol were a mixture of co-morbidities, complications and disease manifestations of spondyloarthropathies, which needed further classification. For example psoriasis and inflammatory bowel disease can be considered comorbidities of spondyloarthritis. It was also discussed that it is important for clinicians to know that inflammatory bowel disease can co-occur with SpA and that the presence of IBD should therefore be noted in people with suspected or confirmed SpA.

The GDG noted that not all complications of SpA are measureable and that they are frequently dependent upon the patient's awareness and the patient reporting any change in their condition to progress to further investigation (e.g. X-ray).

Cardiovascular risk

The consensus opinion of the committee was that they thought it was important to screen for cardiovascular risk factors. The committee further discussed that the evidence presented did not provide sufficient evidence on the occurrence of aortic insufficiency, as this would occur about 20 years into any study. It was noted that there are screening programs in existence for cardiovascular risk factors.. The GDG were of the opinion that this was an important issue that needed to be highlighted due to the increased risk of mortality if nothing was implemented. The GDG were concerned that there is the risk of appointment fatigue if people with SpA are attending multiple appointments for different risk screening associated with |

their SpA

Iritis

The GDG were concerned that people with iritis have a tendency to self-medicate with over the counter medicines. The GDG thought it was important to highlight to people with SpA and clinicians that if a person with SpA has a sore eye then they should seek help from an appropriately qualified health care professional at the earliest opportunity. However, they noted that recommendations on assessment for suspected uveitis had been made elsewhere in the guideline and did not feel that there was need for an additional recommendation.

Osteoporosis

The decision to screen for osteoporosis every two years was made on the basis of clinical experience, supported by the included study which followed up patients at 24 months. The GDG noted that, in practice, some rheumatologists provide annual reviews to people with spondyloarthritis, including osteoporosis risk assessment, while in other areas this is not the case, so developing a recommendation on this complication was intended to reduce variations in care. The resource impact of this recommendation may therefore vary between areas depending on what practices are already in place.

Fracture

People with SpA have a higher risk of fracture than the general population, even if they do not have osteoporosis (e.g. due to spinal fusion in people with axial spondyloarthritis). The GDG noted that it was difficult to identify osteoporosis in people with SpA with standard screening methods, as anterior-posterior DEXA of the lumbar region may overestimate bone density due to thickened spin the lumbar region. People experiencing changes in pain or occurrence of deformity may attribute this to disease progression or flare episodes, and may not present for investigation of possible fracture. People with SpA need to be aware that low impact incidents can lead to fractures.

The GDG discussed how the general osteoporosis guidance may have different thresholds for assessing osteoporosis than those used in people with SpA, as people with SpA are at increased risk of vertebral fracture compared to the general population

8.6.5 Recommendations

32. **Discuss risk factors for cardiovascular comorbidities with all people with spondyloarthritis.**

33. **Consider regular osteoporosis assessments (every 2 years) for people with axial spondyloarthritis. Be aware that bone mineral density measures may be elevated on spinal DEXA due to the presence of syndesmophytes and ligamentous calcification, whereas hip measurements may be more reliable.**

34. **Advise people with axial spondyloarthritis that they may be prone to fractures, and should consult a healthcare professional following falls or physical trauma, particularly in the event of increased musculoskeletal pain.**

8.6.6 Research recommendations

17. What is the optimum approach for identifying and managing osteoporosis and fracture risk in axial spondyloarthritis?

Why this is important

Risks of osteoporosis and fracture are known to be higher in people with axial spondyloarthritis than the general population. However, few studies have looked at whether this higher risk means it is appropriate to adopt a different strategy for identifying and monitoring these conditions in this group (e.g. is more intensive monitoring justified). Prospective cohort studies addressing questions about the frequency of monitoring could help to improve outcomes for this high-risk group.

8.7 Complications of treatments for spondyloarthritis

Review Question 33

- What are the complications associated with treatments for spondyloarthritis?

8.7.1 Evidence review

Patients with spondyloarthritis may be at risk of complications associated with different treatment options.

The aim of this review was to identify the complications associated with treatment for spondyloarthritis so that these can be added to information given to patients and can be monitored for in any regular patient review, managing the risk where appropriate.

Table 38: PICO table – complications of spondyloarthritis

Population	People (aged 16 years and over) with a confirmed diagnosis of spondyloarthritis
Complications	NSAIDs; gastritis, ulcers, bleeding, cardiovascular events (potential risk reduction), renal, hypertension
	Corticosteroids; cataracts, diabetes, osteoporosis, suppressed adrenal gland hormone production, thin skin, easy bruising and slower wound healing, weight gain, (wound) infection, psychosis, hypertension
	Standard DMARDs; myelosuppression, renal toxicity, liver toxicity, skin rash, gastrointestinal disturbance, malignancy, hypertension, haematological toxicity
	Biological DMARDs; infection, immunosuppression, malignancy (especially skin), demyelination, progressive multifocal leukoencephalopathy, depression, skin rash, uveitis
	Intra-articular and soft tissue injections; injections, local steroid effect, skin depigmentation, fat necrosis, tendon rupture
Comparators	Not applicable
Outcomes	Rates of each complication (follow-up of over 1 year)

A single systematic search was conducted for both question 32 (complications of condition) and question 33 (complications of treatments), which identified a total of 13,303 references. Cohort studies were considered to be the optimal study design. Follow up studies of randomised controlled trial populations were also considered for inclusion if the follow up period was of at least one year (so as to distinguish between short-term adverse events and long-term complications. The references were screened on their titles and abstracts and 184 studies were ordered for full text for question 33. 171 studies were excluded as they did not meet the eligibility criteria such as complications not specified in the scope, or studies without a defined follow up period. A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F. An additional 18 studies were identified through rerun searches at the end of the guideline, leading to a total of 31 studies being included.

8.7.1.1 Description of included studies

Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G

8.7.1.1.1 *Biological DMARDs:*

Twenty seven studies were identified which contained information on adverse events occurring to people taking biological DMARDs. Twenty one of these studies were either cohort studies or single-arm extension studies following randomised controlled trials, and 6 were based on data from registries. Fifteen of the studies contained data on people with ankylosing spondylitis, 2 contained data on people with axial spondyloarthritis, 12 contained

data on people with psoriatic arthritis, and 3 contained data on people with undifferentiated spondyloarthritis.

8.7.1.1.2 *Standard DMARD*

Two studies were identified which contained information on adverse events occurring to people taking standard DMARDs, both based on data from registries. One of the studies contained data on people with axial spondyloarthritis and 1 contained data on people with psoriatic arthritis.

8.7.1.1.3 *NSAIDs*

One study was identified which contained information on adverse events occurring to people taking NSAIDs, based on data from a registry. This study contained data on people with ankylosing spondylitis and undifferentiated spondyloarthritis.

8.7.1.1.4 *Corticosteroids*

One study was identified which contained information on adverse events occurring to people taking glucocorticoids, based on data from a registry. This study contained data on people with axial spondyloarthritis.

8.7.1.2 Variations from protocol

The initial protocol for this question included the specification that the rates of each potential complication of treatment should be given at pre-defined time points. After very little evidence was identified in an initial search, this criterion was agreed to be too restrictive by the GDG, who agreed that it should be removed to allow the inclusion of studies where events had been reported across the follow-up period.

8.7.1.3 Minimal clinically important differences

Since the majority of studies identified in this review were non-comparative, no minimal clinically important differences were considered.

8.7.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

8.7.3 Evidence statements

Very low-quality evidence from 27 studies reported rates of adverse events for people with spondyloarthritis taking biological DMARDs, standard DMARDs, NSAIDs and corticosteroids. The adverse events reported were in line with those expected from these medicines, including an elevated rate of malignancy, and in particular skin cancer, compared to general population averages. The event rates were commonly lower than for people with rheumatoid arthritis taking the same medicines. Rates for specific complications are presented the evidence tables in Appendix E.

8.7.4 Evidence to recommendations

| Relative value of different outcomes | The GDG agreed that the outcomes included in the review protocol were important; of these they noted the potential clinical severity of the pre-specified outcomes, demyelination and progressive multifocal leukoencephalopathy. The GDG further discussed the long-term complications of treatment that people with spondyloarthritis who are |

	contemplating initiation or changing of therapy, (in particular biological DMARDs) are often most concerned about. They considered these to be immunosuppression and risk of cancer.
Trade-off between benefits and harms	The GDG discussed the evidence relating to long-term complications of treatment for spondyloarthritis. Overall it was concluded that the evidence presented reflected the known potential complications from long-term biologic therapy and that for people with spondyloarthritis there were no additional complications. The possible exception to this was with psoriatic arthritis. It was discussed that people with psoriatic arthritis who undergo PUVA (Psoralen with UVA) treatment for comorbid psoriasis may have an elevated risk of skin cancer and this should be taken into consideration when advising people about the benefits and risks associated with biological DMARDs.
	The GDG acknowledged that the treatments given to people with spondyloarthritis have known side effects or risks of complications alongside the benefits, as reported in studies of people with other conditions who receive these treatments. The GDG considered the evidence from the presented studies in people with spondyloarthritis.
	The GDG agreed that it is important that people with spondyloarthritis should have access to information about both the benefits and harms over the short- and long-term. The relatively young age at onset or diagnosis for some of those with spondyloarthritis (specifically psoriatic arthritis) was noted and that this may have implications for treatment decision making in relation to possible complications of both their condition and the treatment. The GDG noted that in their experience, people were particularly concerned about cancer risks, and that this may sometimes dissuade people from initiating anti-TNF therapies or standard DMARD therapies who may otherwise have benefitted. However, it was noted that some people with spondyloarthritis are keen to access these therapies as soon as possible. In the GDG's experience an individual's perception of risk may therefore be an important determinant of treatment choice. The GDG noted that there has historically been concern about possible increased malignancy risk in people taking anti-TNFs for other indications, and that this may also apply to this population, though not necessarily at sufficiently high rates to outweigh the potential treatment benefits.
	It was noted that, in the experience of the GDG, monitoring of side effects or complications of treatment may vary between clinical specialties e.g. people with psoriatic arthritis may experience more intense monitoring from dermatology than from rheumatology.
	The GDG noted that people with spondyloarthritis may seek information about benefits and harms of treatment from websites, and that this information may not always provide sufficient context in which to assess the likelihood of harm. They agreed the importance of discussion with people with spondyloarthritis of any concerns and questions they may have about long-term complications.
Economic considerations	No studies were found in the health economic review and no de novo modelling was conducted for this review question. The GDG did not believe that its recommendations would incur any new opportunity costs, as they were confined to providing advice.
Quality of evidence	The GDG agreed that the evidence identified did meet the review protocol and the inclusion of registry data was appropriate in this review question.

	All identified studies were assessed as being of very low quality and most lacked comparison groups for outcome data. It was noted that a number of studies which may provide evidence of shorter-term complications had been omitted from the review, as these did not meet the review criteria of follow up of at least one year duration. The GDG agreed on the GRADE assessment of the relevant outcomes of this evidence as very low.
	The GDG discussed the low rate of depression reported in the single study that contained data on this outcome. It was noted that, although chronic health conditions often have higher rates of comorbid depression, that successful treatment of pain and arthritis symptoms may relieve depression in some people with spondyloarthritis.
Other considerations	The GDG opted not to make specific clinical recommendations relating to the long-term complications of treatment for spondyloarthritis, noting that there was no evidence presented which suggested that the complications from treatment in these conditions were any different from complications from these treatments in other groups. However, they did opt to make a "be aware" recommendation about the potential for higher levels of skin cancer with biological DMARDs. The GDG noted that skin cancer is a well-established complication of biological DMARD use in other populations, and it was therefore felt to be a particularly important to make people with spondyloarthritis aware of this, though the risks are not sufficiently great to negate the benefits of these therapies for managing this group of conditions. They agreed that there was insufficient evidence to suggest a need for a different approach to monitoring and surveillance overall.

8.7.5 Recommendations

35. Be aware there may be a greater risk of skin cancer in people treated with TNF-alpha inhibitors.

9 Information for people with spondyloarthritis

When spondyloarthritis is diagnosed, it is important that people receive information to help them to understand the condition. People need information and support to make decisions about treatments and how to self-manage spondyloarthritis. People also need information about the potential long-term complications of spondyloarthritis and information including effective strategies to manage flare episodes; in particular how and whom to contact within the multidisciplinary team.

It should not be assumed that every person diagnosed with spondyloarthritis will have the same health beliefs, circumstances, preferences, information and education needs; consideration of those factors should be paramount in discussions about what each person wants or needs to manage their condition.

People with spondyloarthritis also need support and reassurance that, as they age and the condition progresses, the support they need will be available when they need it. This might include information about access to specialist services within the NHS, and signposting to external resources. Local and national charities, voluntary organisations and support groups may act as a useful and important adjunct by providing accredited, evidence-based, tailored knowledge and mutual support.

The aim of information provision for people with spondyloarthritis should be to enhance clinical care and provide ongoing support. Information and support should be effective and appropriate, and take into consideration each person's wider needs beyond the initial diagnosis and the treatment pathway.

9.1 Information for people with spondyloarthritis

Review Question 27

- What information on treatment, long-term complications and self-management do young people and adults with spondyloarthritis find useful?

9.1.1 Evidence review

The aim of this review was to assess what information on spondyloarthritis, assessments, diagnosis, treatment options, long-term complications and self-management young people and adults with spondyloarthritis find useful.

The review focussed on identifying studies that fulfilled the conditions specified in Table 39.

Table 39: PICO table for information and education for people with spondyloarthritis

Population	Patients with spondyloarthritis
Interventions	Information for patients (including information on: treatment, long term-complications, self-management)
Comparators	Different formats of information, different content of information, timing of provision of information, delivery setting
Outcomes	Patient and clinician reported outcomes (including: usefulness, accuracy)

For full details of the review protocol please see Appendix C.

Randomised controlled trials were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well.

A systematic search and a hand search of the reference lists of systematic reviews identified 1,900 references. The references were screened on their titles and abstracts and 37 studies were ordered for full text.

Thirty-three studies were excluded as they did not meet the eligibility criteria for reasons such as inappropriate population (e.g. young children, adolescents younger than 16 years of age, or not primarily spondyloarthritis patients) or inappropriate outcome (i.e. no measure of usefulness of information). A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F.

A total of 4 studies was identified for inclusion.

Studies identified

The included studies featured those with populations of adults with spondyloarthritis and those covering adolescents and young adults, whose information needs may be considered slightly different. Evidence tables for included studies can be found in Appendix E, with GRADE profiles reported in Appendix G

9.1.1.1 Description of included studies

The 4 included studies (Cooksey, 2012 Dragoi 2013 Giacomelli 2015 Leung 2009) were carried out in Austria, Hong Kong, Italy and the UK, The study sample size ranged from 105 to 743 and males accounted for between 42% and 81% of participants in the included studies. The age of participants and mean duration of disease ranged from a mean of 50.3 years (SD=12.2) to a mean of 57 years (SD=13) and a mean of 9.8 years (SD=6.9) to a mean of 23 years (SD 14) respectively but neither were reported in 1 study (Giacomelli 2015). Two studies (Dragoi 2013, Leung 2009) included patients with a diagnosis of psoriatic arthritis only, 1 study (Cooksey 2012) included only patients with ankylosing spondylitis while

1 study (Giacomelli 2015) included mixed populations of both ankylosing spondylitis and psoriatic arthritis.

Patients were asked to complete, anonymously and independently, a specifically developed questionnaire during their scheduled rheumatology consultation. There were 60 questions in 14 domains, including those related to information provision (Giacomelli 2015). In Cooksey 2012, participants were asked to complete a postal or online questionnaire about information needs, The questionnaire consisted of open and close-ended questions. The responses to open ended questionnaire items were explored for patterned responses and emerging themes using thematic analysis. In Leung 2009, patients consented to self-administer questionnaires on demographic data, quality of life, adequacy of perceived care, participation in medical decision, satisfaction with health care and specific health care needs.

In Dragoi 2013, an educational needs assessment tool (ENAT self-report questionnaire) was used in a cross-sectional survey to assess the relationship between educational needs, disease activity and function. Patients were asked to complete the tool at a routine visit to the rheumatology outpatient clinic. The questionnaire included 39 items grouped into 7 domains (e.g. managing pain, movement, feelings etc.), each item was rated on a Likert scale of 1 (not important at all) to 4 (extremely important).

9.1.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

9.1.3 Evidence statements

Four studies provided low or very low quality evidence on the information needs of patients with ankylosing spondylitis or psoriatic arthritis and identified the following as themes:

- Summaries on latest research and medications
- Stories and experiences from other AS patients
- Opportunity to ask a doctor questions
- AS networking
- Diagnosis, medication, exercises and how to improve performance of daily activities
- Information on disease
- Advice on exercise
- Use of alternative medicine
- Managing pain
- Arthritis process
- Treatments
- Self-help measures

9.1.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that the findings of the studies summarised in the evidence statement included those domains that they would consider important to those with spondyloarthritis.
	The GDG discussed how changing disease classification for some of the spondyloarthritis conditions may have made some existing information out-of-date and noted the importance of information being current.
	The GDG noted the importance of ensuring that, where information is

	provided or information sources highlighted to people with spondyloarthritis, this needs to be appropriate to the stage of the person's condition, age and other current circumstances for the person at that time. The GDG discussed the need to individualise any discussions with people with spondyloarthritis and the depth of information that will be requested may be different from person to person.
Trade-off between benefits and harms	The GDG agreed the need for information to be available in formats that patients of all ages and stages in their disease could access. They expressed concern that it may be difficult for people with spondyloarthritis to make the distinction between reliable sources of information and those that may not be evidence-based or supported by national organisations or charities. This may be particularly true for information accessed via search engines and social media. They agreed the importance of sign-posting people to reliable sources of information and the need for discussing the importance of using reliable sites with patients. The GDG considered that many of the spondyloarthritis conditions can be difficult for patients to explain as there may not be any symptoms that may be obvious to others. They considered that providing people with help with explaining their condition (e.g. to employers, families) would be an area where support should be provided.
Economic considerations	No economic evidence was presented
Quality of evidence	Overall the GDG noted there was a lack of good quality evidence in this area. They considered that while this evidence was relevant, it may not encompass the needs of all people with spondyloarthritis. The GDG identified that the information needs of younger patients or those with early onset conditions may currently be unmet (e.g. if information is not available in formats considered as accessible or useful). They discussed the different methods of accessing information that may be used and how these may differ across patient groups. They agreed the importance of providing information in a variety of formats e.g. leaflets, social media, apps. The need to ensure that appropriate information is provided on an on-going basis was discussed by the GDG. They noted that this should include information relevant to the stage of a person's condition and information on how to access to specialist services when needed, for example during flare episodes. Accepting the limitations of the evidence presented the GDG considered the available evidence and using their own expertise and clinical experience developed a consensus based recommendation.
Other considerations	The GDG noted that the NICE patient experience guideline would be relevant guidance with which to cross-refer. They additionally opted to make recommendations specific to this population based on their experience. These recommendations included what information should be considered as appropriate to give in relation to flare episodes.

1 ### 9.1.5 Recommendations

2 **Information about spondyloarthritis**

3 **36. Provide people with spondyloarthritis, and their family members or carers (as**
4 **appropriate), with information that is:**

5 • available on an ongoing basis
6 • relevant to the stage of the person's condition

tailored to the person's needs.

For more guidance on providing information to people and discussing their preferences with them, see the NICE guideline on patient experience in adult NHS services.

37. Provide explanations and information about spondyloarthritis. Information should be oral and written, and may include:

- what spondyloarthritis is
- diagnosis and prognosis
- treatment options (pharmacological and non-pharmacological)
- likely symptoms and how they can be managed
- flares and extra-articular symptoms
- self-help options
- research and medicines
- which healthcare professionals will be involved with the person's care and how to get in touch with them.
- local support groups, online forums and national charities, and how to get in touch with them.

Information about disease flares

38. Advise people with spondyloarthritis about the possibility of experiencing flare episodes and extra-articular symptoms.

39. Consider developing a flare management plan that is tailored to the person's individual needs, preferences and circumstances.

40. When discussing any flare management plan, provide information on:

- access to care during flares
- self-care (for example, exercises, stretching and joint protection)
- pain and fatigue management
- medicines
- managing the impact on daily life and ability to work.

9.2 Information and education for flare management in spondyloarthritis

Review Question 28

- What is the effectiveness of information and education in the management of flare episodes?

9.2.1 Evidence review

The aim of this review was to assess the effectiveness of information and education in the management of flare episodes in (axial and peripheral) spondyloarthritis.

The review focussed on identifying studies that fulfilled the conditions specified in Table 40.

Table 40: PICO table for information and education in flare management

Population	Patients with spondyloarthritis
Interventions	Education and information for patients and clinicians on management of flare episodes (including: who to contact, how to self-manage, when to contact, identification of flare episodes)
Comparators	Standard information given to patients, no information
Outcomes	Patient reported outcomes (including: usefulness of information in terms of being able to access care or self-management, number and duration of flare episodes (i.e. poorly controlled disease), number of contacts with HCP, patient satisfaction)

For full details of the review protocol please see Appendix C

Randomised controlled trials were considered to be the highest quality evidence available to answer this question and are graded as high in a GRADE framework if conducted and reported well.

A systematic search and a hand search of the reference lists of systematic reviews identified 1,867 references. The references were screened on their titles and abstracts and 8 studies were ordered for full text.

All studies were excluded as they did not meet the eligibility criteria such as inappropriate study design (not an RCT), inappropriate population (not primarily spondyloarthritis) or inappropriate intervention (education in the context of exercise rather than flare management). A detailed list of excluded studies and reasons for their exclusion is provided in Appendix F.

9.2.1.1 Description of included studies

No studies met the inclusion criteria for this review

9.2.2 Health economic evidence

A literature search was conducted jointly for all review questions in this guideline by applying standard health economic filters to a clinical search for spondyloarthritis (see Appendix C). A total of 9,970 references was retrieved, of which none were retained for this review question. Health economic modelling was not prioritised for this review question.

9.2.3 Evidence statements

No evidence was identified

9.2.4 Evidence to recommendations

Relative value of different outcomes	The GDG agreed that it is important to consider the effectiveness of information and education in the management of flare episodes to ensure that people with spondyloarthritis can recognise their own symptoms and when to seek healthcare advice and from which services. The GDG noted that there can be variation in current practice as to what information people with spondyloarthritis access during flare episodes, with some initially accessing primary care services and some going directly to specialist services.
Trade-off between benefits and harms	The GDG considered that the difficulties with defining flare make providing recommendations in this area difficult. They noted that individuals will experience flare episodes in different ways with a variety of symptoms. This makes the provision of a generic management strategy challenging. Triggers of flare and changes in symptoms may be difficult to distinguish from each other, which may affect the advice given on flare management.
Economic considerations	No economic evidence was presented or discussed
Quality of evidence	No RCT evidence was identified from the searches that considered the effectiveness of information and education in the management of flare episodes as applicable to this review question. The GDG agreed the usefulness of the specialist team discussing flare with people with spondyloarthritis and agreeing individualised plans on how to manage flare episodes. These could include initial self-management and whether access to primary care or specialist care would be most appropriate. The importance of recognising flare and managing it to prevent any further possible disease progression or complication was also noted. The GDG further noted the need to discuss with the patient their flare episodes and where it may be possible to identify other contributing factors that may influence the reoccurrence of flare episodes.
Other considerations	The GDG opted not to make a specific recommendation relating to effectiveness of information provision during flare management, but did make recommendations about general aspects of flare-related information (see Q27)

9.2.5 Recommendations

No recommendation was drafted

9.2.6 Research recommendation

18. What approaches to signposting people with spondyloarthritis to appropriate services for managing their flares are found most useful by people with spondyloarthritis?

Why this is important

Being provided with appropriate information about flares is important for people with spondyloarthritis, but there is a lack of evidence about the most appropriate ways to ensure people have access to this information, and whether this differences between different subgroups of the population. Qualitative studies of preferences for information in people with spondyloarthritis who have experience of flares would enable the optimisation of support services for people at risk of having flares.

19. What is the effectiveness and cost-effectiveness of information provision in reducing the incidence and severity of flare episodes?

Providing structured information about flares may help to reduce their incidence and severity, but there is a cost attached to providing these services. Well conducted randomised controlled trials would help to show whether there are any benefits from such an approach, and if these benefits are sufficiently large to justify the cost of providing this information prospectively.

10 Glossary

10.1 Abbreviations

Abbreviations used in this guideline	
AS	ankylosing spondylitis
ASAS	Assessment of SpondyloArthritis international Society
ASQoL	Ankylosing Spondylitis Quality of Life Questionnaire
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BASG	Bath Ankylosing Spondylitis Global Score
BASMI	Bath Ankylosing Spondylitis Metrology Index
CRP	C-reactive protein
DEXA	Dual-energy X-ray absorbiometry, a types of scan for assessing bone mineral density
DMARDs	disease modifying anti-rheumatic drugs (standard or biological)
Ent-SpA	enteropathic spondyloarthritis
EQ-5D	EuroQol five dimensions questionnaire (quality of life measure)
ESR	erythrocyte sedimentation rate
HAQ	Health Assessment Questionnaire
HLA-B27	Human leukocyte antigen-B27
ICER	incremental cost-effectiveness ratio
JIA	juvenile idiopathic arthritis
MRI	magnetic resonance imaging
nr-axSpA	non-radiographic axial spondyloarthritis
NSAIDs	non-steroidal anti-inflammatory drugs
PASI	Psoriasis Area and Severity Index
PDUS	power Doppler ultrasound
PsA	psoriatic arthritis
Ps/PsO	psoriasis
QALY	quality-adjusted-life year
ReA	reactive arthritis
SpA	spondyloarthritis
SPC	summary of product characteristics (for medicines)

1 **10.2 Glossary**

Glossary of terms and abbreviations used in this guideline	
ankylosing spondylitis	An inflammatory condition predominantly involving the spine from the sacroiliac joints upwards. It causes progressive restriction of spinal movement due to calcification of spinal ligaments and bony change and fusion of the spinal joints. It may affect peripheral joints and be associated with inflammation in other areas including the eye and gut
axial spondyloarthritis	An inflammatory condition primarily affecting the spinal joints, including the sacroiliac joints. The term includes ankylosing spondylitis and non-radiographic axial spondyloarthritis. Axial joints include hips and joints around the sternum. Peripheral joints and entheses may also be involved
Biological DMARDs	Disease modifying anti-rheumatic drugs that target specific proteins which are involved in the initiation and perpetuation of inflammation in inflammatory arthritides
chronic back pain	defined in this guideline as lower back pain of 3 months or more duration
C-reactive protein	a biomarker of inflammation, measured by a blood test
dactylitis	inflammation of a finger or toe characteristically resulting in a sausage appearance of the affected digit
enteropathic spondyloarthritis	spondyloarthritis associated with Crohn's disease or ulcerative colitis
enthesitis	inflammation of an enthesis (site at which a tendon or ligament attaches to bone)
erythrocyte sedimentation rate	a measure of inflammation, measured by a blood test
flare	acute exacerbation of symptoms/disease activity above the level usually experienced
Human leukocyte antigen-B27.	A genetic marker associated with spondyloarthritis, detected by a blood test for particular proteins on white blood cells
hydrotherapy	water-based physiotherapy, usually in a dedicated pool at a higher temperature than a standard swimming pool
inflammatory arthritis	group of conditions, including spondyloarthritis, where there is inflammation of the joints
inflammatory back pain	back pain characteristic of axial spondyloarthritis, often chronic and associated with stiffness, which may improve with exercise or movement and NSAIDs but be worse at rest
juvenile idiopathic arthritis	inflammatory arthritis condition, with onset in childhood or adolescence (aged under 16)
monoarthritis	arthritis affecting a single joint
morning stiffness	symptom particularly experienced in inflammatory arthritis where stiffness of the joints and/or spine is experienced on waking for at least 30 minutes but may improve through the day
non-radiographic axial spondyloarthritis	axial spondyloarthritis where sacroiliitis or spinal inflammatory disease is not detectable on plain X-ray
non-selective NSAIDs	NSAIDs which do not specifically target the COX-2 enzyme
oligoarthritis	arthritis affecting 4 or fewer joints
peripheral spondyloarthritis	An inflammatory condition primarily affecting the peripheral joints (non-axial joints), which is distinct from conditions like Rheumatoid arthritis. Psoriatic arthritis, enteropathic arthritis and reactive arthritis are usually predominantly peripheral, though may have some axial involvement
polyarthritis	arthritis affecting more than 4 joints

Glossary of terms and abbreviations used in this guideline	
psoriasis	A chronic skin condition which causes thickened scaly and silvery patches. It can also affect the scalp and the nails
psoriatic arthritis	peripheral spondyloarthritis associated with psoriasis (N.B. In some cases, psoriatic arthritis will occur before psoriasis develops or there is only a family history of it)
range of motion exercise	exercises designed to help joints move through the full range of movement
reactive arthritis	arthritis arising from an abnormal immune response triggered by some infections. Bacterial triggers within the scope of this guideline are Campylobacter, Chlamydia, Salmonella, Shigella, and Yersinia
sacroiliitis	inflammation of the sacroiliac joint at the base of the spine
scintigraphy	radiological imaging technique using radioisotopes
selective NSAIDs	NSAIDs which specifically target the COX-2 enzymes involved in inflammation
spondyloarthritis	Group of conditions characterised by inflammatory arthritis, excluding rheumatoid arthritis. The group includes ankylosing spondyloarthritis, non-radiographic axial spondyloarthritis, psoriatic arthritis, enteropathic arthritis, reactive arthritis and undifferentiated spondyloarthritis. Historically referred to as seronegative arthropathies.
standard DMARD	Disease modifying anti-rheumatic drugs that alter the underlying disease process in a less specific to biological DMARDs
syndesmophyte	An abnormal growth of bone arising in the spinal ligaments which occurs as part of the inflammatory process in spondyloarthritis
synovitis	inflammation of synovium, a tissue that lines many joints and tendon sheathes
undifferentiated spondyloarthritis	Spondyloarthritis which does not fit into any of the other diagnostic categories
uveitis, (acute anterior)	inflammation of the uveal tract (middle layer of the eye), which includes the front of the iris

1

Appendices

Appendix A: <Insert first appendix heading here>