National Institute for Health and Care Excellence

Consultation

Rheumatoid arthritis in adults: diagnosis and management

Evidence review A Ultrasound for diagnosis

NICE guideline CG79

Diagnostic evidence review

January 2018

Consultation

This evidence review was developed by the National Guideline Centre

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Contents

1 Ultrasound for diagnosis of rheumatoid arthritis 6

1.1 Review question: In adults with suspected inflammatory arthritis (including rheumatoid arthritis), what is the added value of ultrasound in the diagnosis of rheumatoid arthritis? 6

1.2 Introduction 6

1.3 PICO table 6

1.4 Clinical evidence 7

1.4.1 Included studies 7

1.4.2 Excluded studies 8

1.4.3 Summary of clinical studies included in the evidence review 8

1.4.4 Quality assessment of clinical studies included in the evidence review 11

1.5 Economic evidence 14

1.5.1 Included studies 14

1.5.2 Excluded studies 14

1.5.3 Unit costs 14

1.6 Resource costs 15

1.7 Evidence statements 15

1.7.1 Clinical evidence statements 15

1.7.2 Health economic evidence statements 16

1.8 Recommendations 17

1.8.1 Research recommendations 17

1.9 Rationale and impact 17

1.9.1 Why the committee made the recommendations 17

1.9.2 Why we need recommendations on this topic 17

1.10 The committee's discussion of the evidence 17

1.10.1 Interpreting the evidence 17

1.10.2 Cost effectiveness and resource use 20

1.10.3 Other factors the committee took into account 20

Appendices 27

Appendix A: Review protocols 27

Appendix B: Literature search strategies 32

B.1 Clinical search literature search strategy 32

B.2 Health Economics literature search strategy 34

Appendix C: Clinical evidence selection 38

Appendix D: Clinical evidence tables 39

Appendix E: Coupled sensitivity and specificity forest plots 49

Appendix F: Health economic evidence selection 51

Appendix G: Health economic evidence tables 53

Appendix H: Excluded studies 54
H.1 Excluded clinical studies 54
H.2 Excluded health economic studies 55
Appendix I: Research recommendations 56
I.1 Ultrasound in cases of diagnostic uncertainty 56

1 Ultrasound for diagnosis of rheumatoid arthritis

1.1 Review question: In adults with suspected inflammatory arthritis (including rheumatoid arthritis), what is the added value of ultrasound in the diagnosis of rheumatoid arthritis?

1.2 Introduction

Most people with rheumatoid arthritis (RA) have definite synovitis on clinical assessment, but there is sometimes uncertainty about the diagnosis when there is no definite synovitis. This can lead to a delay in starting treatment, which could affect prognosis.

Use of ultrasound with clinical assessment may be more effective than clinical assessment alone at identifying synovitis and thereby diagnosing rheumatoid arthritis. Ultrasound may also allow healthcare professionals to be more confident about ruling out a diagnosis of rheumatoid arthritis.

1.3 PICO table

For full details, see the review protocol in appendix A.

Table 1: PICO characteristics of clinical effectivess review

Population	Adults with suspected inflammatory arthritis (including rheumatoid arthritis)
Interventions	• Clinical assessment plus ultrasound
Comparison	• Clinical assessment without ultrasound
Outcomes	CRITICAL – CLINICAL EFFECTIVENESS OUTCOMES • Disease Activity Score (continuous) at 12 months • Quality of life at 12 months • Function at 12 months IMPORTANT – PROCESS OUTCOMES • Definitive clinical diagnosis (dichotomous) at time of testing • Change/reclassification of diagnosis (dichotomous) by end of the study (or post ultrasound) • Change in management (dichotomous) at time of testing • Prescribed DMARDs (dichotomous) at time of testing • Require repeat testing / additional testing (dichotomous) at time of testing
Study design	Randomised controlled trials (RCTs) Systematic Review / Network Meta-Analysis of RCTs

Table 2: PICO characteristics of diagnostic accurat review

Population	Adults with suspected inflammatory arthritis (including rheumatoid arthritis)
Target condition	Rheumatoid arthritis
Index test	Ultrasound plus clinical assessment of any joints
Reference	Clinical diagnosis of rheumatoid arthritis

standard	Clinical diagnosis may be made either 'on the spot' or at a later date (for example, 3-12 months following testing). Greater weight will be placed on data where the diagnosis is made after at least 3 months follow up.
Statistical measures and outcomes	• CRITICAL – DIAGNOSTIC ACCURACY OUTCOMES • Sensitivity • Specificity • Positive predictive value • Negative predictive value • Area under the curve (AUC) IMPORTANT – PROCESS OUTCOMES • Definitive clinical diagnosis (dichotomous) at time of testing • Change/reclassification of diagnosis (dichotomous) by end of the study (or post ultrasound) • Change in management (dichotomous) at time of testing • Prescribed DMARDs (dichotomous) at time of testing • Require repeat testing / additional testing (dichotomous) at time of testing
Study design	Diagnostic accuracy studies

This review sought to investigate clinical assessment plus ultrasound in 2 stages. Firstly the review sought out randomised cotrolled trials comparing diagnosis with clinical assessment combined with ultrasound versus diagnosis via clinical assessment alone. The outcomes would give a comparison of the clinical effectiveness of the diagnostic methods.

The second strage assessed the diagnostic accuracy of clinical assessment plus ultrasound using diagnosis via clinical assessment in the future as the gold standard. In the absence of a gold standard method for diagnosing RA, future assessment was agreed by the committeeas more reliable as the signs of synovitis will be much more pronounced from a clinical assessment perspective.

Sensitivity was considered the most critical outcome. This is because failing to diagnose people who have rheumatoid arthritis may delay the initiation of DMARD treatment and reduce the likelihood of the person achieving long-term remission or low disease activity. A minimum threshold of 90% sensitivity was set for recommending the test.

In addition, a number of process outcomes were considered important for both sections of the review. These were definitive clinical diagnosis, change or reclassification of diagnosis, change in planned management, prescription of DMARDs, and requirement for repeat or additional testing.

1.4 Clinical evidence

1.4.1 Included studies

A search was conducted for randomised controlled trials, diagnostic accuracy studies and systematic reviews of these study types assessing the clinical effectiveness or diagnostic accuracy of clinical assessment of any joints with ultrasound in people with suspected inflammatory arthritis.

Four diagnostic accuracy studies were included in the review;[10,17,27,30] these are summarised in Table 3 below. All 4 studies evaluated the diagnostic accuracy of clinical

assessment with ultrasound and one of the studies evaluated the change or reclassification of diagnosis following ultrasound.

Evidence from these studies is summarised in the clinical evidence summary below (Table 4 and Table 5).

See also the study selection flow chart in appendix C, sensitivity and specificity forest plot in appendix E, and study evidence tables in appendix D.

1.4.2 Excluded studies

See the excluded studies list in appendix H.

1.4.3 Summary of clinical studies included in the evidence review

Table 3: Summary of diagnostic accuracy studies included in the evidence review

Study	Population	Target condition	Tests	Reference standard	Comments
Filer 2011[10]	People with clinically apparent synovitis of at least 1 joint and inflammatory joint symptoms for ≤3 months. N=58	Rheumatoid arthritis	Index tests: 1: Gray-scale US combined with 1987 ACR criteria. 2. Power Doppler US combined with 1987 ACR criteria.	Diagnosis according to 1987 ACR criteria:18 month follow-up	Ultrasound evaluated 38 joints in in hands, feet, wrists, elbow, shoulder, knee and ankle. Shoulder, elbow, knee and ankle ultrasound variables discarded from analysis due to low specificity for RA. Unclear how US combined with criteria. Very serious risk of bias due to no details of how participants were selected and no specification of how the ACR criteria were supplemented with ultrasound results The study was assessed to be applicable and direct evidence.
Ji 2017[17]	People with arthritic complaints and 1 tender joint and/or swollen joint in the hand with inflammatory	Rheumatoid arthritis	Index tests: 1. 2010 ACR/EULAR score combined with US GS total score	1987 ACR criteria after at least 1 year follow-up (median: 15 months)	Ultrasound assessment of 22 joints in the hands and wrists. Very serious risk of bias due to

Study	Population	Target condition	Tests	Reference standard	Comments
	joint symptoms Additionally: negative ACPA and no bone erosions on x-ray. N=94		2. 2010 ACR/EULAR score combined with US PD total score 3. 2010 ACR/EULAR score combined with US synovitis joint count		unclear reporting of index test analysis and selection of participants not indicated to be consecutive. The study was assessed to be applicable and direct evidence.
Nakagomi 2013[27]	Consecutive people with musculoskeletal problems for ≤3 years with possible diagnosis of RA. People with no clinically swollen joints were not excluded in order to include people with subclinical synovitis. N=109	Rheumatoid arthritis	Index tests 1. 2010 ACR/EULAR classification criteria but joint distribution was replaced with US GS synovitis score of ≥1 2. 2010 ACR/EULAR classification criteria but joint distribution was replaced with US GS synovitis GS score of ≥2 or PD score ≥1	2010 ACR/EULAR criteria at baseline (no follow-up)	Ultrasound assessment of 38 joints in hands, feet, wrists, elbow, shoulder, knee and ankle. No diagnostic accuracy data. The study was assessed to be applicable and direct evidence. Low risk of bias for change /reclassification of diagnosis. Serious risk of bias for diagnostic outcomes due to reference standard test happening at baseline.
Navalho 2013[30]	Consecutive people with untreated clinically apparent synovial swelling. Involvement of at least 1 joint of wrists or hands. N=45	Rheumatoid arthritis	Index test ACR/EULAR 2010 classification criteria where US joint and tendon counts replaced clinical joint counts	1987 ACR criteria at 12 months follow-up.	Ultrasound procedure was limited to the wrists and hands. Low risk of bias. The study was assessed to be applicable and direct evidence.

1 See appendix D for full evidence tables.

1

2

1

1.4.4 2 Quality assessment of clinical studies included in the evidence review

3 **Table 4: Clinical evidence summary: ultrasound plus clinical assessment**

Index Test (Threshold)	Number of studies	n	Quality	Specificity % & Sensitivity % (95% CI)	Positive predictive value (PPV) & negative predictive value (NPV)	AUC (95% CI)
Gray-scale ultrasound combined with 1987 ACR criteria	1	58	VERY LOW[1,2] due to risk of bias and imprecision	Sensitivity: 93% (77% - 99%) Specificity: 66% (46% - 82%)	PPV: 73% NPV: 91%	AUC: 0.793
Power Doppler ultrasound combined with 1987 ACR criteria	1	58	VERY LOW[1,2] due to risk of bias and imprecision	Sensitivity: 86% (68% - 96%) Specificity: 76% (56% - 90%)	PPV: 78% NPV: 85%	AUC: 0.810
2010 ACR/EULAR score or ≥2 joints with synovitis in the hands	1	94	LOW[1] due to risk of bias	Sensitivity: 86%		
2010 ACR/EULAR score combined with GS total score	1	94	LOW[1] due to risk of bias			AUC: 0.864
2010 ACR/EULAR score combined with PD total score	1	94	LOW[1] due to risk of bias			AUC: 0.869
2010 ACR/EULAR score combined with synovitis joint count	1	94	LOW[1] due to risk of bias			AUC: 0.872
ACR/EULAR 2010 classification criteria with US	1	45	HIGH			AUC: 0.948 (0.836-0.992)
2010 ACR/EULAR classification criteria but joint distribution was replaced with US GS synovitis score of ≥1	1	109	LOW[1,2] due to risk of bias and imprecision	Sensitivity: 82% (67% - 93%) Specificity: 75% (64% - 85%)	PPV: 66% NPV: 88%	

Index Test (Threshold)	Number of studies	n	Quality	Specificity % & Sensitivity % (95% CI)	Positive predictive value (PPV) & negative predictive value (NPV)	AUC (95% CI)
ACR/EULAR classification criteria but joint distribution was replaced with US GS synovitis GS score of ≥2 or PD score ≥1	1	109	LOW[1,2] due to risk of bias and imprecision	Sensitivity: 57% (41% - 73%) Specificity: 90% (80% - 96%)	PPV: 77% NPV: 78%	

The assessment of the evidence quality was conducted with emphasis on sensitivity as this was identified by the committee as the primary measure in guiding decision-making

1. *Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.*
2. *Imprecision was assessed based on inspection of the confidence region for sensitivity in the diagnostic analysis. The evidence was downgraded by 1 increment when there was a 20-40% range of the confidence interval around the point estimate, and downgraded by 2 increments when there was a range of >40%*

Table 5: Clinical evidence summary of process outcome: change/reclassification of diagnosis

Comparison	Number of studies	n	Quality	Preliminary classification	Alteration to preliminary classifications	Comments
Index test: 2010 ACR/EULAR classification criteria but joint distribution was replaced with US GS synovitis score of ≥1 Comparator test: 2010 ACR/EULAR classification criteria	1	109	HIGH	Preliminary diagnosis: Index test: RA: 50, not-RA: 59 Comparator test: RA: 40, not-RA: 69	17 people reclassified as having RA after index test. 7 People reclassified as not having RA after index test.	Comparator test undertaken first and followed by index test on the same day.
Index test: 2010 ACR/EULAR classification criteria but joint distribution was replaced with US GS synovitis GS	1	109	HIGH	Preliminary diagnosis: Index test: RA: 30, not-RA: 79 Comparator test: RA: 40, not-RA: 69	7 people reclassified as having RA after index test. 17 People reclassified as not having RA after index test	Comparator test undertaken first and followed by index test on the same day.

Comparison	Number of studies	n	Quality	Preliminary classification	Alteration to preliminary classifications	Comments
score of ≥2 or PD score ≥1 Comparator test: 2010 ACR/EULAR classification criteria						

1

2

1.5 Economic evidence

1.5.1 Included studies

No relevant health economic studies were identified.

1.5.2 Excluded studies

No health economic studies that were relevant to this question were excluded due to assessment of limited applicability or methodological limitations.

See also the health economic study selection flow chart in appendix F.

1.5.3 Unit costs

The unit costs of rheumatology appointments and of unbundled diagnostic ultrasound imaging are provided below for guidance.

Table 6: Cost of outpatient rheumatology appointments

Currency Code	Currency Description	No. of attendances	National Average Unit Cost
Consultant led			
WF01A	Non-Admitted Face to Face Attendance, Follow-Up	1,223,574	£137
WF01B	Non-Admitted Face to Face Attendance, First	311,626	£220
WF02A	Multi-professional Non-Admitted Face to Face Attendance, Follow-Up	7,357	£218
WF02B	Multi-professional Non-Admitted Face to Face Attendance, First	4,219	£246
Non-consultant led			
WF01A	Non-Admitted Face to Face Attendance, Follow-Up	250,578	£87
WF01B	Non-Admitted Face to Face Attendance, First	59,478	£146
WF02A	Multi-professional Non-Admitted Face to Face Attendance, Follow-Up	928	£106
WF02B	Multi-professional Non-Admitted Face to Face Attendance, First	366	£114

Source: NHS Reference costs, 2015-2016[3]

Table 7: Cost of ultrasound

Department Description(a)	Currency Code	Currency Description	No. of examinations	National Average Unit Cost
Direct Access	RD40Z	Ultrasound Scan with duration of less than 20 minutes, without contrast	1,905,598	£51
Direct Access	RD41Z	Ultrasound Scan with duration of less than 20 minutes, with contrast	43,644	£39
Direct Access	RD42Z	Ultrasound Scan with duration of 20 minutes and over, without contrast	463,721	£60

Department Description(a)	Currency Code	Currency Description	No. of examinations	National Average Unit Cost
Direct Access	RD43Z	Ultrasound Scan with duration of 20 minutes and over, with contrast	23,462	£52
Direct Access	RD44Z	Ultrasound Scan, Mobile or Intraoperative Procedures, with duration of less than 20 minutes	31,126	£42
Direct Access	RD45Z	Ultrasound Scan, Mobile or Intraoperative Procedures, with duration of 20 to 40 minutes	22,770	£99
Outpatient	RD40Z	Ultrasound Scan with duration of less than 20 minutes, without contrast	1,993,859	£55
Outpatient	RD41Z	Ultrasound Scan with duration of less than 20 minutes, with contrast	48,731	£52
Outpatient	RD42Z	Ultrasound Scan with duration of 20 minutes and over, without contrast	519,666	£66
Outpatient	RD43Z	Ultrasound Scan with duration of 20 minutes and over, with contrast	20,377	£66
Outpatient	RD44Z	Ultrasound Scan, Mobile or Intraoperative Procedures, with duration of less than 20 minutes	28,758	£55
Outpatient	RD45Z	Ultrasound Scan, Mobile or Intraoperative Procedures, with duration of 20 to 40 minutes	64,212	£89
Other	RD40Z	Ultrasound Scan with duration of less than 20 minutes, without contrast	18,468	£56
Other	RD42Z	Ultrasound Scan with duration of 20 minutes and over, without contrast	3,556	£88
Weighted average				**£55**

Source: NHS Reference costs, 2015-2016[3]
(a) Direct access services are provided independently of an admission or outpatient attendance because a patient is referred by a GP for a test or self-refers.

1.6 Resource costs

The recommendations made in this review are not expected to have a substantial impact on resources.

1.7 Evidence statements

1.7.1 Clinical evidence statements

The evidence on diagnostic accuracy was inconsistent within studies, dependent on how ultrasound was integrated into the diagnostic process, and also across studies. The sensitivity and specificity of the test ranged from 93% and 66% to 57% and 90% (2 studies, low to very low quality, n=167). Other measures of accuracy were AUC which varied from 0.79 to 0.95 (3 studies, very low to high quality, n=197), PPV which varied from 66% to 78%

(2 studies, low to very low quality, n=167), and NPV which varied from 78% to 91% (2 studies, low to very low quality, n=167).

Evidence on change or reclassification of diagnosis reported that the use of ultrasound changed diagnoses, but without follow-up it is not known whether the reclassification was correct (1 study, high quality evidence, n=109).No evidence was available for any of the clinical effectiveness outcomes.

1.7.2 Health economic evidence statements

No relevant economic evaluations were identified.

1.8 Recommendations

[No recommendation]

1.8.1 Research recommendations

A.RR1. What is clinical and cost effectiveness of using ultrasound in addition to clinical assessment when there is uncertainty about the diagnosis in adults with suspected RA?

See also the rationale in appendix I.

1.9 Rationale and impact

1.9.1 Why the committee made the recommendations

Ultrasound is not used widely in diagnosing RA, but use is increasing and depends on the clinic and the rheumatologist. Evidence was inconsistent and too limited for the committee to make any recommendation for or against its use in diagnosis. The committee noted that the studies generally included only people with clinically definite synovitis and agreed that ultrasound may be more useful when there is uncertainty about the diagnosis after clinical assessment. They decided to make a research recommendation to inform future guidance on who (if anyone) should have ultrasound to aid diagnosis.

1.9.2 Why we need recommendations on this topic

Most people with rheumatoid arthritis have definite synovitis on clinical assessment, but there is sometimes uncertainty about the diagnosis when there is no definite synovitis. This can lead to a delay in starting treatment, which could affect prognosis.

Use of ultrasound with clinical assessment may be more effective than clinical assessment alone at identifying synovitis and thereby diagnosing rheumatoid arthritis. Ultrasound may also allow healthcare professionals to be more confident about ruling out a diagnosis of rheumatoid arthritis.

1.10 The committee's discussion of the evidence

1.10.1 Interpreting the evidence

1.10.1.1 The outcomes that matter most

The review was split into 2 components. The clinical effectiveness aspect of the review aimed to establish whether use of ultrasound in diagnosis improves patient outcomes. For this part of the review, the committee agreed that the most critical outcome was disease activity, as the overall benefit of early diagnosis and early treatment should be captured in reduced disease activity scores. The next 2 critical outcomes were quality of life and function, which have a complementary role in terms of describing the overall impact of the disease on a person's life.

In the section of the review focussing on determining the diagnostic accuracy of ultrasound in addition to clinical assessment in diagnosing rheumatoid arthritis, sensitivity was considered the most critical outcome. This is because failing to diagnose people who have rheumatoid arthritis may delay the initiation of DMARD treatment and reduce the likelihood of the person achieving long-term remission or low disease activity. A minimum threshold of 90%

sensitivity was set for recommending the test. Specificity was also considered critical because DMARDs have adverse events and cost implications and so should not be used unnecessarily; however, it was agreed that priority would be placed on sensitivity. Other accuracy statistics considered important were positive and negative predictive values; and area under the curve (AUC), which provides an overall summary of the test performance.

In addition, a number of process outcomes were considered important for both sections of the review. These were definitive clinical diagnosis, change or reclassification of diagnosis, change in planned management, prescription of DMARDs, and requirement for repeat or additional testing.

No evidence was identified for any of the clinical effectiveness outcomes, or any of the process outcomes other than change or reclassification of diagnosis.

1.10.1.2 The quality of the evidence

No randomised controlled trials (RCTs) were identified that compared a diagnostic strategy using ultrasound and clinical assessment with a diagnostic strategy of clinical assessment alone to establish the impact on patient outcomes.

Four prospective cohort studies were included that evaluated the diagnostic accuracy of ultrasound plus clinical assessment for rheumatoid arthritis. Three of these studies used the index test in the participants to make a preliminary diagnosis of rheumatoid arthritis and followed participants up 12 or more months later to confirm or refute the diagnosis using the classification criteria for rheumatoid arthritis (see section 1.10.1.3 below).

The committee agreed that the studies using the classification criteria complemented with the ultrasound data, compared to a reference standard of the classification criteria applied after follow up of at least 12 months, were the most reliable way of assessing the diagnostic accuracy of the addition of ultrasound. One of the studies did not involve long-term follow-up and instead the index test was compared to the classification criteria as the reference standard at a single point in time. This was assessed to be at serious risk of bias for this reason, and the committee placed less weight on this data.

None of the data were able to be meta-analysed due to the differences between the index tests and the outcomes reported. The quality of the diagnostic accuracy evidence was assessed per index test and ranged from high quality for 1 test (from 1 study), low quality for 6 tests (from 2 studies) and very low quality for 2 tests (from 1 study). Most of the evidence was assessed to be at serious or very serious risk of bias, often due to unclear methods of participant selection and poor reporting of index test analysis (for example, in 1 study it was unclear how the ultrasound variables were integrated into the index test).

Most of the studies reported only AUC statistics rather than sensitivity and specificity data. AUC is an overarching measure of accuracy of a test and does not give an indication of the trade-off between sensitivity and specificity of the test, and therefore was less informative to the committee for decision making. Where sensitivity and specificity were not reported, it was not possible to assess evidence quality fully, as imprecision could not be assessed, which further reduced the committee's confidence in the AUC evidence. For those studies that did report sensitivity data, the confidence intervals around the estimates of sensitivity were very wide, so the committee also considered this evidence highly uncertain.

1.10.1.3 Benefits and harms

The evidence for the use of ultrasound in the diagnosis of rheumatoid arthritis was highly heterogeneous. The included studies enrolled different populations, used different index tests and study designs, and reported accuracy data in different ways. The committee also noted that some of the results were conflicting.

It was noted that the AUC data were inconsistent as data from 1 study suggested that overall the use of ultrasound reduced diagnostic accuracy compared to clinical assessment alone, whereas the other 2 studies reporting AUC showed an increase in diagnostic accuracy with ultrasound, compared to clinical assessment alone.

The committee agreed that the data on change or reclassification of diagnosis were interesting, as they showed that the addition of ultrasound data did impact preliminary diagnostic decisions; 22% of diagnoses were altered based on the additional ultrasound variable. However, as there was no longer term follow up of the study participants, it was impossible to know whether the reclassification was correct. For that reason, this evidence was not given substantial weight in the committee's deliberations.

The committee also placed little weight on the sensitivity and specificity data from the study that did not involve long-term follow-up, consistent with the approach agreed in the review protocol. In this study, the index test (classification criteria plus ultrasound) was compared to the classification criteria alone as the reference standard, applied at the same point in time. The committee noted that this design meant that any added benefit of ultrasound over the existing classification criteria would not be captured, as the reference standard is assumed to be 100% accurate. The lack of long-term follow-up also rendered the results unreliable.

Most weight was placed on the evidence from sensitivity and specificity of the index test with the reference standard applied after 18 months follow-up, as the committee considered this study design to be the most useful in answering the clinical question about the added value of ultrasound. This evidence suggested that using ultrasound in the diagnosis of rheumatoid arthritis may improve sensitivity compared to clinical assessment alone (by reducing the number of true cases missed) at the expense of specificity (by increasing the number of cases incorrectly diagnosed as having rheumatoid arthritis). The committee agreed it was reasonable to expect that diagnosis using ultrasound would miss fewer people with rheumatoid arthritis than diagnosis without ultrasound, potentially through the detection of subclinical synovitis caused by rheumatoid arthritis that may otherwise have been overlooked. It was also not unexpected that the use of ultrasound may diagnose a proportion of people with rheumatoid arthritis incorrectly, as some of the ultrasound-detected synovitis may have had a non-rheumatoid arthritis cause. However, the committee agreed that even this evidence was not overly persuasive, as it was based on a single small, low quality study.

Overall, the committee considered that the evidence from the 4 small heterogeneous studies was too limited and of insufficient quality to support any recommendation about the use of ultrasound in diagnosis of rheumatoid arthritis. It was agreed that from this limited evidence and consensus opinion, ultrasound was unlikely to be a useful tool in the diagnosis of everyone with suspected rheumatoid arthritis but the evidence is not sufficiently strong to make a definitive recommendation to this effect.

Crucially, the committee discussed that most of the studies (including the one the committee considered most informative) enrolled people with clinically definite synovitis. The committee was of the view that ultrasound was most likely to be of most benefit in diagnosing a subgroup of people with suspected rheumatoid arthritis *without* clinically definite synovitis. It was considered that this mismatch between the broader populations enrolled in the trials and the narrower subgroup of potential interest, may mean that the included studies were unable to capture the potential benefit of ultrasound in the subgroup. The committee agreed to make a research recommendation to establish the added value of ultrasound in diagnosing rheumatoid arthritis where there is diagnostic uncertainty following clinical assessment. This subgroup may include people with symptoms of rheumatoid arthritis but without clinically definite synovitis.

1.10.21 Cost effectiveness and resource use

No health economic evidence was identified. The unit costs of ultrasound and a rheumatology outpatient appointment were presented to the committee to aid the consideration of cost-effectiveness. The committee reviewed the unit cost of the ultrasound (£55) and felt that this cost was likely to reflect the cost of ultrasound undertaken in a radiology department rather than in a rheumatology department. They noted that when ultrasound is used for diagnostic purposes, it is often done by the rheumatologist within the rheumatology department rather than referred to radiology in order to avoid any delays in diagnosis due to referral wait time. The committee thought that the unit cost of ultrasound conducted by a rheumatologist was likely to be greater than £55.

The committee noted that MRI is sometimes used in current practice for the purpose of diagnosis and that ultrasound could be a cheaper alternative in those circumstances.

The committee noted that based on the clinical evidence reviewed, there was insufficient evidence to make a recommendation for the use of ultrasound in diagnosis and agreed that a research recommendation was needed. They discussed that in a subset of people, in whom a diagnosis is not possible on clinical assessment alone; there may be a benefit of ultrasound to complement clinical assessment. In these people, the committee suggested that the cost of ultrasound might be offset by the benefits of a prompt diagnosis or early discharge if RA is not diagnosed.

The lack of recommendation is unlikely to have a resource impact, as current practice will continue.

1.10.22 Other factors the committee took into account

Currently, the diagnosis of rheumatoid arthritis relies primarily on clinical examination and judgment with supportive investigations. Classification criteria developed by the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) are widely used as eligibility criteria in clinical trials. The 2010 ACR/EULAR criteria involve the assessment of the number and type of involved joints, serology (RF and ACPA status), inflammatory markers (ESR and CRP), and duration of symptoms, to calculate a score out of 10, with a score of at least 6 to be classified as having definite rheumatoid arthritis. The earlier 1987 ACR classification criteria also included factors such as morning stiffness and radiographic changes, and required the presence of at least 4 of 7 criteria for classification as rheumatoid arthritis. The classification criteria are not designed to be used as diagnostic criteria in clinical practice. However, diagnosis in clinical practice does draw on factors included in the criteria, but without necessarily tallying a total score. The committee agreed that the application of the classification criteria as a reference standard applied after follow up was the best way of assessing the diagnostic utility of ultrasound plus clinical assessment, in the absence of a ‘gold standard’ test for rheumatoid arthritis.

The committee noted that in some cases people with rheumatoid arthritis are reluctant to accept their diagnosis and commence treatment. Ultrasound may help improve patient outcomes in these circumstances by enabling clinicians to show people objective evidence of their joint inflammation and thereby encourage them to commence appropriate therapy. Further research should help to clarify the circumstances where ultrasound assessment may be clinically and cost effective in diagnosing rheumatoid arthritis.

The time taken to conduct the scan was reported in 2 of the studies. One study indicated that the US examination took 50-60 minutes and the second study indicated each scan took at least 15 minutes. The former study evaluated hands, wrists, shoulder, elbow, knee and ankle joints while the latter study was limited to joints in the hands and wrists. The committee indicated time to complete a scan would be faster with a sonographer who would be direct and focused, while a rheumatologist would be utilising the session for a broader purpose and will interact with the person being scanned in a more investigative fashion. The

rheumatologist would be utilising their expertise to investigate the possible diagnosis and make on-the-spot decisions based on the clinical and ultrasound tests. Additionally, rheumatologist-conducted ultrasound is likely to be undertaken more quickly for a person with suspected rheumatoid arthritis, and this could result in faster diagnosis and treatment.

The committee also acknowledged that ultrasound is not the only additional test used in diagnosing rheumatoid arthritis; MRI is also sometimes used, at greater additional cost. This review did not consider the relative costs and benefits of MRI and ultrasound.

1

References

1. Agrawal S, Bhagat SS, Dasgupta B. Improvement in diagnosis and management of musculoskeletal conditions with one-stop clinic-based ultrasonography. Modern Rheumatology. 2009; 19(1):53-56

2. Aydin SZ, Gunal EK, Ozata M, Keskin H, Ozturk AB, Emery P et al. Six-joint ultrasound in rheumatoid arthritis: a feasible approach for implementing ultrasound in remission. Clinical and Experimental Rheumatology. 2017; 35(5):853-856

3. Barlow JH, Barefoot J. Group education for people with arthritis. Patient Education and Counseling. 1996; 27(3):257-267

4. Botar-Jid C, Bolboaca S, Fodor D, Bocsa C, Tamas MM, Micu M et al. Gray scale and power Doppler ultrasonography in evaluation of early rheumatoid arthritis. Medical Ultrasonography. 2010; 12(4):300-305

5. Broll M, Albrecht K, Tarner I, Muller-Ladner U, Strunk J. Sensitivity and specificity of ultrasonography and low-field magnetic resonance imaging for diagnosing arthritis. Clinical and Experimental Rheumatology. 2012; 30(4):543-547

6. Chaiamnuay S, Lopez-Ben R, Alarcon GS. Ultrasound of target joints for the evaluation of possible inflammatory arthropathy: associated clinical factors and diagnostic accuracy. Clinical and Experimental Rheumatology. 2008; 26(5):875-880

7. D'Agostino MA, Haavardsholm EA, van der Laken CJ. Diagnosis and management of rheumatoid arthritis; What is the current role of established and new imaging techniques in clinical practice? Best Practice and Research in Clinical Rheumatology. 2016; 30(4):586-607

8. D'Agostino MA, Terslev L, Wakefield R, Ostergaard M, Balint P, Naredo E et al. Novel algorithms for the pragmatic use of ultrasound in the management of patients with rheumatoid arthritis: From diagnosis to remission. Annals of the Rheumatic Diseases. 2016; 75(11):1902-1908

9. El Miedany Y, Youssef S, Mehanna AN, El Gaafary M. Development of a scoring system for assessment of outcome of early undifferentiated inflammatory synovitis. Joint, Bone, Spine: Revue du Rhumatisme. 2008; 75(2):155-162

10. Filer A, de Pablo P, Allen G, Nightingale P, Jordan A, Jobanputra P et al. Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. Annals of the Rheumatic Diseases. 2011; 70(3):500-507

11. Freeston JE, Wakefield RJ, Conaghan PG, Hensor EM, Stewart SP, Emery P. A diagnostic algorithm for persistence of very early inflammatory arthritis: the utility of power Doppler ultrasound when added to conventional assessment tools. Annals of the Rheumatic Diseases. 2010; 69(2):417-419

12. Ha AS, Plotkin B. Ultrasound techniques in rheumatoid arthritis. Current Radiology Reports. 2016; 4(9):1-7

13. Hirata A, Ogura T, Hayashi N, Takenaka S, Ito H, Mizushina K et al. Concordance of patient-reported joint symptoms, physician-examined arthritic signs, and ultrasound-detected synovitis in rheumatoid arthritis. Arthritis Care & Research. 2017; 69(6):801-806

14. Hmamouchi I, Bahiri R, Srifi N, Aktaou S, Abouqal R, Hajjaj-Hassouni N. A comparison of ultrasound and clinical examination in the detection of flexor tenosynovitis in early arthritis. BMC Musculoskeletal Disorders. 2011; 12:91

15. Horton SC, Tan AL, Wakefield RJ, Freeston JE, Buch MH, Emery P. Ultrasound-detectable grey scale synovitis predicts future fulfilment of the 2010 ACR/EULAR RA classification criteria in patients with new-onset undifferentiated arthritis. RMD Open. 2017; 3(1):e000394

16. Hurnakova J, Hulejova H, Zavada J, Komarc M, Hanova P, Klein M et al. Serum Calprotectin Discriminates Subclinical Disease Activity from Ultrasound-Defined Remission in Patients with Rheumatoid Arthritis in Clinical Remission. PloS One. 2016; 11(11):e0165498

17. Ji L, Deng X, Geng Y, Song Z, Zhang Z. The additional benefit of ultrasonography to 2010 ACR/EULAR classification criteria when diagnosing rheumatoid arthritis in the absence of anti-cyclic citrullinated peptide antibodies. Clinical Rheumatology. 2017; 36(2):261-267

18. Kamel SR, Sadek HA, Mohamed FA, Osman HM. Role of ultrasound disease activity score in assessing inflammatory disease activity in rheumatoid arthritis patients. Egyptian Rheumatologist. 2017; Epublication

19. Kawashiri SY, Suzuki T, Okada A, Yamasaki S, Tamai M, Nakamura H et al. Musculoskeletal ultrasonography assists the diagnostic performance of the 2010 classification criteria for rheumatoid arthritis. Modern Rheumatology. 2013; 23(1):36-43

20. Komarova EB. Correlation between joint ultrasound data and clinical, laboratory and instrumental indices in patients with rheumatoid arthritis. Sovremennye Tehnologii v Medicine. 2017; 9(1):92-96

21. Lage-Hansen PR, Lindegaard H, Chrysidis S, Terslev L. The role of ultrasound in diagnosing rheumatoid arthritis, what do we know? An updated review. Rheumatology International. 2017; 37(2):179-187

22. Lai KL. New applications of ultrasound for the diagnosis and treatment of rheumatoid arthritis. Journal of Medical Ultrasound. 2016; 24(4):129-131

23. Mankia K, Emery P. Preclinical rheumatoid arthritis: progress toward prevention. Arthritis & Rheumatology. 2016; 68(4):779-788

24. Mathew AJ, Danda D, Conaghan PG. MRI and ultrasound in rheumatoid arthritis. Current Opinion in Rheumatology. 2016; 28(3):323-329

25. Millot F, Clavel G, Etchepare F, Gandjbakhch F, Grados F, Saraux A et al. Musculoskeletal ultrasonography in healthy subjects and ultrasound criteria for early arthritis (the ESPOIR cohort). Journal of Rheumatology. 2011; 38(4):613-620

26. Minowa K, Ogasawara M, Murayama G, Gorai M, Yamada Y, Nemoto T et al. Predictive grade of ultrasound synovitis for diagnosing rheumatoid arthritis in clinical practice and the possible difference between patients with and without seropositivity. Modern Rheumatology. 2016; 26(2):188-193

27. Nakagomi D, Ikeda K, Okubo A, Iwamoto T, Sanayama Y, Takahashi K et al. Ultrasound can improve the accuracy of the 2010 American College of Rheumatology/European League against rheumatism classification criteria for rheumatoid arthritis to predict the requirement for methotrexate treatment. Arthritis & Rheumatism. 2013; 65(4):890-898

28. Naredo E, Iagnocco A. One year in review: ultrasound in arthritis. Clinical and Experimental Rheumatology. 2016; 34(1):1-10

29. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

30. Navalho M, Resende C, Rodrigues AM, Pereira da Silva JA, Fonseca JE, Campos J et al. Bilateral evaluation of the hand and wrist in untreated early inflammatory arthritis: a comparative study of ultrasonography and magnetic resonance imaging. Journal of Rheumatology. 2013; 40(8):1282-1292

31. Ohrndorf S, Backhaus M. Pro musculoskeletal ultrasonography in rheumatoid arthritis. Clinical and Experimental Rheumatology. 2015; 33(4 Suppl 92):S50-53

32. Ozgul A, Yasar E, Arslan N, Balaban B, Taskaynatan MA, Tezel K et al. The comparison of ultrasonographic and scintigraphic findings of early arthritis in revealing rheumatoid arthritis according to criteria of American College of Rheumatology. Rheumatology International. 2009; 29(7):765-768

33. Plaza M, Nowakowska-Plaza A, Pracon G, Sudol-Szopinska I. Role of ultrasonography in the diagnosis of rheumatic diseases in light of ACR/EULAR guidelines. Journal of Ultrasonography. 2016; 16(64):55-64

34. Ponikowska M, Wiland P. The initial ultrasonographic examination of hands and feet joints in patients with early rheumatoid arthritis. Reumatologia. 2015; 53(4):179-185

35. Pratt AG, Lorenzi AR, Wilson G, Platt PN, Isaacs JD. Predicting persistent inflammatory arthritis amongst early arthritis clinic patients in the UK: is musculoskeletal ultrasound required? Arthritis Research & Therapy. 2013; 15:R118

36. Rakieh C, Nam JL, Hunt L, Hensor EM, Das S, Bissell LA et al. Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study. Annals of the Rheumatic Diseases. 2015; 74(9):1659-1666

37. Rezaei H, Torp-Pedersen S, af Klint E, Backheden M, Kisten Y, Gyori N et al. Diagnostic utility of musculoskeletal ultrasound in patients with suspected arthritis--a probabilistic approach. Arthritis Research & Therapy. 2014; 16:448

38. Rizzo G, Raffeiner B, Coran A, Ciprian L, Fiocco U, Botsios C et al. Pixel-based approach to assess contrast-enhanced ultrasound kinetics parameters for differential diagnosis of rheumatoid arthritis. Journal of Medical Imaging. 2015; 2(3):034503

39. Salaffi F, Ciapetti A, Gasparini S, Carotti M, Filippucci E, Grassi W. A clinical prediction rule combining routine assessment and power Doppler ultrasonography for predicting progression to rheumatoid arthritis from early-onset undifferentiated arthritis. Clinical and Experimental Rheumatology. 2010; 28(5):686-694

40. Schmidt WA. Value of sonography in diagnosis of rheumatoid arthritis. Lancet. 2001; 357(9262):1056-1057

41. Sizova L. Diagnostic value of antibodies to modified citrullinated vimentin in early rheumatoid arthritis. Human Immunology. 2012; 73(4):389-392

42. Sizova L. Imaging methods of joint damage in early rheumatoid arthritis. Current Rheumatology Reviews. 2015; 11(1):64-67

43. Takase-Minegishi K, Horita N, Kobayashi K, Yoshimi R, Kirino Y, Ohno S et al. Diagnostic test accuracy of ultrasound for synovitis in rheumatoid arthritis: systematic review and meta-analysis. Rheumatology. 2017; Epublication

44. Tamas MM, Rednic N, Felea I, Rednic S. Ultrasound assessment for the rapid classification of early arthritis patients. Journal of Investigative Medicine. 2013; 61(8):1184-1191

45. Valor L, Martinez-Estupinan L, Janta I, Nieto JC, Ovalles-Bonilla JG, Gonzalez-Fernandez C et al. Ultrasound-detected joint inflammation and B cell count: related variables for rituximab-treated RA patients? Rheumatology International. 2016; 36(6):793-797

46. van de Stadt LA, Bos WH, Meursinge Reynders M, Wieringa H, Turkstra F, van der Laken CJ et al. The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study. Arthritis Research & Therapy. 2010; 12:R98

47. van der Ven M, Kuijper TM, Gerards AH, Tchetverikov I, Weel AE, van Zeben J et al. No clear association between ultrasound remission and health status in rheumatoid arthritis patients in clinical remission. Rheumatology. 2017; 56(8):1276-1281

48. Zhao CY, Jiang YX, Li JC, Xu ZH, Zhang Q, Su N et al. Role of Contrast-enhanced Ultrasound in the Evaluation of Inflammatory Arthritis. Chinese Medical Journal. 2017; 130(14):1722-1730

49. Zufferey P, Rebell C, Benaim C, Ziswiler HR, Dumusc A, So A. Ultrasound can be useful to predict an evolution towards rheumatoid arthritis in patients with inflammatory polyarthralgia without anticitrullinated antibodies. Joint, Bone, Spine: Revue du Rhumatisme. 2017; 84(3):299-303

1 Appendices

2 Appendix A: Review protocols

3 **Table 8: Review protocol: Ultrasound for diagnosis of rheumatoid arthritis**

ID	Field	Content
I	Review question	In adults with suspected inflammatory arthritis (including rheumatoid arthritis), what is the added value of ultrasound in the diagnosis of rheumatoid arthritis?
II	Type of review question	Combined diagnostic accuracy and clinical effectiveness review A review of health economic evidence related to the same review question was conducted in parallel with this review. For details see the health economic review protocol for this NICE guideline.
III	Objective of the review	In people with suspected inflammatory arthritis, synovitis is often detectable on clinical assessment. However, in some people with suspected inflammatory arthritis, synovitis is subclinical and this can make diagnosis difficult. The aim of this review is to determine the added value of using ultrasound to assist in the diagnosis of rheumatoid arthritis in patients with suspected inflammatory arthritis.
IV	Eligibility criteria – population / disease / condition / issue / domain	Population: Adults with suspected inflammatory arthritis (including rheumatoid arthritis). Studies in a narrower subgroup of this population will still be included. Target condition: Rheumatoid arthritis
V	Eligibility criteria – intervention(s) / exposure(s) / prognostic factor(s)	Intervention or index test: Ultrasound plus clinical assessment of any joints Ultrasound assessment should be performed by an appropriately trained healthcare professional.
VI	Eligibility criteria – comparator(s) / control or reference (gold) standard	Comparator: Clinical assessment of any joints without an ultrasound element Reference standard: Clinical diagnosis of rheumatoid arthritis. Clinical diagnosis may be made either 'on the spot' or at a later date (for example, 3-12 months following testing). Greater weight will be placed on data where the diagnosis is made after at least 3 months follow up.
VII	Outcomes and prioritisation	CRITICAL – CLINICAL EFFECTIVENESS OUTCOMES • Disease Activity Score (continuous) at 12 months • Quality of life (for example, EQ5D, SF-36, RA Quality of Life instrument, patient global assessment as per OMERACT method) (continuous) at 12 months • Function (for example, Health Assessment Questionnaire, activities of daily living) (continuous) at 12 months. Clinical effectiveness outcome data must be recorded least 6 months after testing. If multiple time points, take closest time point to 12 months. CRITICAL – DIAGNOSTIC ACCURACY OUTCOMES • Sensitivity • Specificity

ID	Field	Content
		• Positive predictive value • Negative predictive value • Area under the curve (AUC). The focus of the accuracy review will be on sensitivity, with a minimum threshold of 90% set for recommending the test. IMPORTANT – PROCESS OUTCOMES • Definitive clinical diagnosis (dichotomous) at time of testing • Change/reclassification of diagnosis (dichotomous) by end of the study (or post ultrasound) • Change in management (dichotomous) at time of testing • Prescribed DMARDs (dichotomous) at time of testing • Require repeat testing / additional testing (dichotomous) at time of testing. Risk association data will not be extracted.
VIII	Eligibility criteria – study design	RCTs Prospective cohort studies Systematic reviews of the above
IX	Other inclusion exclusion criteria	None
X	Proposed sensitivity / subgroup analysis, or meta-regression	None
XI	Selection process – duplicate screening / selection / analysis	A sample of at least 10% of the abstract lists will be double-sifted by a senior research fellow and discrepancies rectified, with committee input where consensus cannot be reached, for more information please see the separate Methods report for this guideline.
XII	Data management (software)	Endnote will be used for bibliographies, citations, sifting and reference management
XIII	Information sources – databases and dates	Clinical search databases: The databases to be searched are Medline, Embase and the Cochrane Library. Date limits for search: None Language: English Health economics search databases: Medline, Embase, NHSEED and HTA Date limits for search: Medline and Embase from 2014 NHSEED and HTA from 2001 Language: English
XIV	Identify if an update	This review is not an update.
XV	Author contacts	https://www.nice.org.uk/guidance/indevelopment/gid-ng10014
XVI	Highlight if	For details, please see section 4.5 of Developing NICE guidelines: the

ID	Field	Content
	amendment to previous protocol	manual.
XVII	Search strategy – for one database	For details, please see appendix B
XVIII	Data collection process – forms / duplicate	A standardised evidence table format will be used, and published as appendix D of the evidence report.
XIX	Data items – define all variables to be collected	For details, please see evidence tables in Appendix D (clinical evidence tables) or H (health economic evidence tables).
XX	Methods for assessing bias at outcome / study level	Diagnostic study checklist (QUADAS 2 tool) will be utilised for quality assessment of diagnostic accuracy outcomes and process outcomes. The risk of bias across all available evidence will be evaluated using a modified GRADE approach.
XXI	Criteria for quantitative synthesis	For details, please see section 6.4 of Developing NICE guidelines: the manual.
XXII	Methods for quantitative analysis – combining studies and exploring (in)consistency	For details, please see the separate Methods report for this guideline.
XXIII	Meta-bias assessment – publication bias, selective reporting bias	For details, please see section 6.2 of Developing NICE guidelines: the manual.
XXIV	Confidence in cumulative evidence	For details, please see sections 6.4 and 9.1 of Developing NICE guidelines: the manual.
XXV	Rationale / context – what is known	For details, please see the introduction to the evidence review.
XXVI	Describe contributions of authors and guarantor	A multidisciplinary committee (https://www.nice.org.uk/guidance/indevelopment/gid-ng10014/documents) developed the evidence review. The committee was convened by the National Guideline Centre (NGC) and chaired by Stephen Ward in line with section 3 of Developing NICE guidelines: the manual. Staff from NGC undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate, and drafted the evidence review in collaboration with the committee. For details, please see Developing NICE guidelines: the manual.
XXVII	Sources of funding / support	NGC is funded by NICE and hosted by the Royal College of Physicians.
XXVIII	Name of sponsor	NGC is funded by NICE and hosted by the Royal College of Physicians.
XXI	Roles of	NICE funds NGC to develop guidelines for those working in the NHS,

ID	Field	Content
X	sponsor	public health and social care in England.
XXX	PROSPERO registration number	Not registered

1

2 **Table 9: Health economic review protocol**

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	Populations, interventions and comparators must be as specified in the clinical review protocol above. Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) Unpublished reports will not be considered unless submitted as part of a call for evidence. Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter – see appendix B below.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2001, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).29 Inclusion and exclusion criteria If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile. If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile. If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. Where there is discretion The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below. The health economist will be guided by the following hierarchies. Setting:

Review question	All questions – health economic evidence
	UK NHS (most applicable). OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). OECD countries with predominantly private health insurance systems (for example, Switzerland). Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. Health economic study type: Cost–utility analysis (most applicable). Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). Comparative cost analysis. Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. Year of analysis: The more recent the study, the more applicable it will be. Studies published in 2001 or later but that depend on unit costs and resource data entirely or predominantly from before 2001 will be rated as 'Not applicable'. Studies published before 2001 will be excluded before being assessed for applicability and methodological limitations. Quality and relevance of effectiveness data used in the health economic analysis: The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

1
2

Appendix B: Literature search strategies

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual 2014, updated 2017. https://www.nice.org.uk/guidance/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869

For more detailed information, please see the Methodology Review.

B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies for interventions as these concepts may not be well described in title, abstract or indexes and therefore difficult to retrieve. Search filters were applied to the search where appropriate.

Table 10: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (Ovid)	1946 – 09 October 2017	Exclusions
Embase (Ovid)	1974 – 09 October 2017	Exclusions
The Cochrane Library (Wiley)	Cochrane Reviews to 2017 Issue 10 of 12 CENTRAL to 2017 Issue 9 of 12 DARE, and NHSEED to 2015 Issue 2 of 4 HTA to 2016 Issue 4 of 4	None

Medline (Ovid) search terms

1.	exp Arthritis, Rheumatoid/
2.	(rheumatoid adj2 (arthritis or arthrosis)).ti,ab.
3.	(caplan* adj2 syndrome).ti,ab.
4.	(felty* adj2 syndrome).ti,ab.
5.	(rheumatoid adj2 factor).ti,ab.
6.	((inflammatory or idiopathic) adj2 arthritis).ti,ab.
7.	"inflammatory polyarthritis".ti,ab.
8.	or/1-7
9.	limit 8 to English language
10.	letter/
11.	editorial/
12.	news/
13.	exp historical article/
14.	Anecdotes as Topic/
15.	comment/

16.	case report/
17.	(letter or comment*).ti.
18.	or/10-17
19.	randomized controlled trial/ or random*.ti,ab.
20.	18 not 19
21.	animals/ not humans/
22.	Animals, Laboratory/
23.	exp Animal Experimentation/
24.	exp Models, Animal/
25.	exp Rodentia/
26.	(rat or rats or mouse or mice).ti.
27.	or/20-26
28.	9 not 27
29.	exp Ultrasonography/
30.	(ultrasound* or ultrason* or echograph* or echotomograph* or doppler).ti,ab.
31.	29 or 30
32.	28 and 31

1 **Embase (Ovid) search terms**

1.	exp *rheumatoid arthritis/
2.	(rheumatoid adj2 (arthritis or arthrosis)).ti,ab.
3.	(caplan* adj2 syndrome).ti,ab.
4.	(felty* adj2 syndrome).ti,ab.
5.	(rheumatoid adj2 factor).ti,ab.
6.	((inflammatory or idiopathic) adj2 arthritis).ti,ab.
7.	"inflammatory polyarthritis".ti,ab.
8.	or/1-7
9.	limit 8 to English language
10.	letter.pt. or letter/
11.	note.pt.
12.	editorial.pt.
13.	case report/ or case study/
14.	(letter or comment*).ti.
15.	or/10-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animal/ not human/
19.	nonhuman/
20.	exp Animal Experiment/
21.	exp Experimental Animal/
22.	animal model/
23.	exp Rodent/
24.	(rat or rats or mouse or mice).ti.
25.	or/17-24
26.	9 not 25
27.	exp *echography/

28.	(ultrasound* or ultrason* or echograph* or echotomograph* or doppler).ti,ab.
29.	27 or 28
30.	26 and 29

Cochrane Library (Wiley) search terms

#1.	[mh "Arthritis, Rheumatoid"]
#2.	(rheumatoid near/2 (arthritis or arthrosis)):ti,ab
#3.	(caplan* near/2 syndrome):ti,ab
#4.	(felty* near/2 syndrome):ti,ab
#5.	(rheumatoid near/2 factor):ti,ab
#6.	((inflammatory or idiopathic) near/2 arthritis):ti,ab
#7.	inflammatory polyarthritis:ti,ab
#8.	(or #1-#7)
#9.	[mh Ultrasonography]
#10.	(ultrasound* or ultrason* or echograph* or echotomograph* or doppler):ti,ab
#11.	#9 or #10
#12.	#8 and #11

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to rheumatoid arthritis population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA) with no date restrictions. NHS EED and HTA databases are hosted by the Centre for Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics studies.

Table 11: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	2014 – 06 October 2017	Exclusions Health economics studies
Embase	2014– 06 October 2017	Exclusions Health economics studies
Centre for Research and Dissemination (CRD)	HTA - 2001 – 06 October 2017 NHSEED - 2001 – 31 March 2015	None

Medline (Ovid) search terms

1.	exp Arthritis, Rheumatoid/
2.	(rheumatoid adj2 (arthritis or arthrosis)).ti,ab.
3.	(caplan* adj2 syndrome).ti,ab.
4.	(felty* adj2 syndrome).ti,ab.
5.	(rheumatoid adj2 factor).ti,ab.
6.	((inflammatory or idiopathic) adj2 arthritis).ti,ab.
7.	"inflammatory polyarthritis".ti,ab.
8.	or/1-7
9.	limit 8 to English language

10.	letter/
11.	editorial/
12.	news/
13.	exp historical article/
14.	Anecdotes as Topic/
15.	comment/
16.	case report/
17.	(letter or comment*).ti.
18.	or/10-17
19.	randomized controlled trial/ or random*.ti,ab.
20.	18 not 19
21.	animals/ not humans/
22.	Animals, Laboratory/
23.	exp animal experiment/
24.	exp animal model/
25.	exp Rodentia/
26.	(rat or rats or mouse or mice).ti.
27.	or/20-26
28.	9 not 27
29.	Economics/
30.	Value of life/
31.	exp "Costs and Cost Analysis"/
32.	exp Economics, Hospital/
33.	exp Economics, Medical/
34.	Economics, Nursing/
35.	Economics, Pharmaceutical/
36.	exp "Fees and Charges"/
37.	exp Budgets/
38.	budget*.ti,ab.
39.	cost*.ti.
40.	(economic* or pharmaco?economic*).ti.
41.	(price* or pricing*).ti,ab.
42.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
43.	(financ* or fee or fees).ti,ab.
44.	(value adj2 (money or monetary)).ti,ab.
45.	or/29-44
46.	exp models, economic/
47.	*Models, Theoretical/
48.	*Models, Organizational/
49.	markov chains/
50.	monte carlo method/
51.	exp Decision Theory/
52.	(markov* or monte carlo).ti,ab.
53.	econom* model*.ti,ab.

54.	(decision* adj2 (tree* or analy* or model*)).ti,ab.
55.	or/46-54
56.	28 and (45 or 55)

1 **Embase (Ovid) search terms**

1.	exp *rheumatoid arthritis/
2.	(rheumatoid adj2 (arthritis or arthrosis)).ti,ab.
3.	(caplan* adj2 syndrome).ti,ab.
4.	(felty* adj2 syndrome).ti,ab.
5.	(rheumatoid adj2 factor).ti,ab.
6.	((inflammatory or idiopathic) adj2 arthritis).ti,ab.
7.	"inflammatory polyarthritis".ti,ab.
8.	or/1-7
9.	limit 8 to English language
10.	letter.pt. or letter/
11.	note.pt.
12.	editorial.pt.
13.	case report/ or case study/
14.	(letter or comment*).ti.
15.	or/10-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animal/ not human/
19.	nonhuman/
20.	exp Animal Experiment/
21.	exp Experimental Animal/
22.	animal model/
23.	exp Rodent/
24.	(rat or rats or mouse or mice).ti.
25.	or/17-24
26.	9 not 25
27.	statistical model/
28.	exp economic aspect/
29.	27 and 28
30.	*theoretical model/
31.	*nonbiological model/
32.	stochastic model/
33.	decision theory/
34.	decision tree/
35.	monte carlo method/
36.	(markov* or monte carlo).ti,ab.

37.	econom* model*.ti,ab.
38.	(decision* adj2 (tree* or analy* or model*)).ti,ab.
39.	or/29-38
40.	*health economics/
41.	exp *economic evaluation/
42.	exp *health care cost/
43.	exp *fee/
44.	budget/
45.	funding/
46.	budget*.ti,ab.
47.	cost*.ti.
48.	(economic* or pharmaco?economic*).ti.
49.	(price* or pricing*).ti,ab.
50.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
51.	(financ* or fee or fees).ti,ab.
52.	(value adj2 (money or monetary)).ti,ab.
53.	or/40-52
54.	26 and (39 or 53)

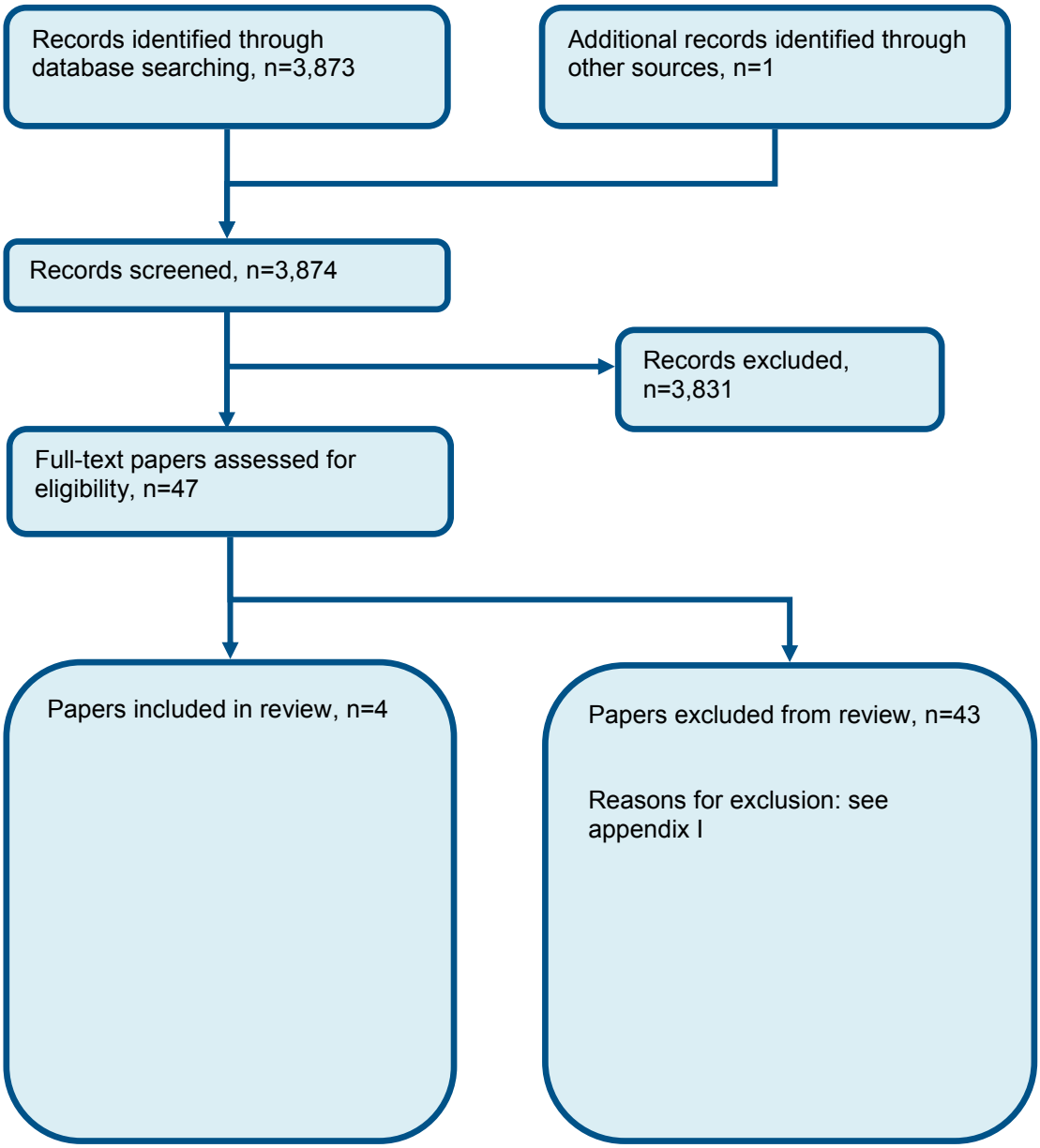
1 **NHS EED and HTA (CRD) search terms**

#1.	MeSH DESCRIPTOR Arthritis, Rheumatoid EXPLODE ALL TREES
#2.	((rheumatoid adj2 (arthritis or arthrosis)))
#3.	((caplan* adj2 syndrome))
#4.	((felty* adj2 syndrome))
#5.	((rheumatoid adj2 factor))
#6.	(((inflammatory or idiopathic) adj2 arthritis))
#7.	("inflammatory polyarthritis")
#8.	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7

2
3

Appendix C: Clinical evidence selection

Figure 1: Flow chart of clinical study selection for the review of ultrasound for diagnosis



Appendix D: Clinical evidence tables

Reference	Filer 2011[10]
Study type	Prospective cohort study
Study methodology	Data source: participant data of cohort study Recruitment: Unclear how and when the participants were recruited. People with clinically apparent synovitis.
Number of patients	n = 58
Patient characteristics	Characteristics reported separately for people diagnosed with RA, diagnosed with non-RA persistent disease, and a resolving group (not clearly defined). Diagnosis by the reference standard. Age, median (range): RA: 63: (19-82), non-RA persistent disease: 45 (18-83), resolving group: 40 (23-75). Gender, male or female (%): RA: 16 (55%) male and 13 (45%) female, non-RA persistent disease: 4 (31%) male and 9 (69%) female, resolving group: 6 (38%) male and 10 (63%) female. Other relevant characteristics: Treatment at baseline: NSAIDs use 41 (68%) Unclear what treatment was received during follow-up. Morning stiffness in minutes, median (range): RA: 120 (30-360), non-RA persistent disease: 60 (0-240), resolving group: 53 (0-240) RF positive, n (%): RA: 15 (52%), non-RA persistent disease: 2 (15%), resolving group:0 (0%) ACPA positive, n (%): RA: 14 (48%), non-RA persistent disease: 0 (0%), resolving group: 0 (0%) ESR mm/h, median (range): RA: 25 (0-104), non-RA persistent disease: 24 (4-87), resolving group: 22 (0-102) CRP MG/L, median (range): RA: 15 (0-102), non-RA persistent disease: 16 (0-83), resolving group: 16 (0-244) SJC of 66, median (range): RA: 8 (1-28), non-RA persistent disease: 2 (1-13), resolving group: 2 (1-7) TJC (of 68) median (range): RA: 9 (0-41), non-RA persistent disease: 3 (0-19), resolving group: 3 (1-10) Presence of erosions, n (%): RA: 11 (38%), non-RA persistent disease: 2 (15%), resolving group: 1 (6%) Family origin: Unclear

Reference	Filer 2011[10]
	Setting: US assessment in radiology suite Country: UK Inclusion criteria: People with clinically apparent synovitis of at least 1 joint and inflammatory joint symptoms (inflammatory joint pain and/or swelling and/or morning stiffness) for 3 months or less. Exclusion criteria: none detailed
Target condition(s)	Rheumatoid arthritis
Index test(s) and reference standard	US assessment Unclear who undertook the ultrasound assessment. Scanner: Siemens Acuson Antares and multi-frequency linear array transducer. Examinations took between 50 and 60 minutes. Person undertaking US assessment was said to be blinded and participants asked not to discuss their symptoms. Undertaken within 24 hours of clinical assessment. Systemic multi-planar gray-scale and power Doppler US examination. Based on EULAR reference scans. Gray-scale synovitis assessment on 0-3 scale and power Doppler positivity and erosion defined according to consensus definitions. Synovial hyperaemia measured by PD and graded 0-3. Index tests 1: Gray-scale US combined with ACR 1987(4/7) criteria. Unclear how US combined with criteria 2. Power Doppler US combined with 1987 ACR (4/7) criteria. Unclear how US combined with criteria Comparator (non US) test 3. 1987 ACR criteria (4/7 clinical) Reference standard Ra diagnosis according to 1987 ACR criteria:18 month follow-up
Statistical measures	Index text: Gray-scale US combined with 1987 ACR (4/7) criteria. Sensitivity (95% CI): 0.93 (0.77-0.99) Specificity (95% CI): 0.655 (0.46-0.82) PPV: 0.73 NPV: 0.91 AUC: 0.793

Reference	Filer 2011[10]
	Index text: Power Doppler US combined with 1987 ACR (4/7) criteria Sensitivity (95% CI): 0.86 (0.68-0.96) Specificity (95% CI): 0.76 (0.56-0.90) PPV: 0.78 NPV: 0.85 AUC: 0.810 Comparator (non US) test: 1987 ACR (4/7 clinical) Sensitivity (95% CI): 0.79 (0.60-0.92) Specificity (95% CI): 0.90 (0.73-0.98) PPV: 0.89 NPV: 0.81 AUC:0.845
Source of funding	Ultrasound equipment funded by Arthritis Research UK and the Rheumatology Research Group is a member of the EU AutoCure Consortium.
Limitations	Risk of bias: very serious risk of bias due to no details of how participants were selected and no specification on how the criteria were supplemented with ultrasound Indirectness: the study was assessed to be applicable and direct evidence.
Comments	

1

Reference	Ji 2017[17]
Study type	Prospective cohort study
Study methodology	Data source: participant data of cohort study Recruitment: Outpatients who had arthritic complaints and visited the Department of Rheumatology and Clinical Immunology at Peking University First Hospital between January 2012 and October 2014 were screened.
Number of patients	n = 94 (29 classified as RA after 1 year)
Patient characteristics	Characteristics reported separately for people with a clinical diagnosis classification of RA after 1 year, and a classified as “non-RA” after 1 year.

Reference	Ji 2017[17]
	Age, mean (SD): RA: 57 (13), non-RA: 51 (17) Gender: female (%): RA: 16 (55%), non-RA: 32 (49%) Other relevant characteristics: Treatment: DMARDs initiated immediately for those diagnosed with RA at baseline. NSAIDs prescribed for those where RA not confirmed and symptoms requiring relief. SJC, median (IQR): RA: 4 (8). Non-RA: 1 (4) TJC, median (IQR): RA: 10 (11). Non-RA: 5 (8) RF positive, n (%): RA: 4 (14%), non-RA: 7 (11%) ESR mm/h, mean (SD): RA: 42 (33), non-RA: 38 (33) CRP MG/L, median (IQR): RA: 16 (22), 10 (27) Ethnicity: Not detailed Setting: Hospital Country: China Inclusion criteria: Outpatients who had arthritic complaints and visited the Department of Rheumatology and Clinical Immunology at Peking University First Hospital. At least 1 tender and/or swollen hand joints with inflammatory joint symptoms (inflammatory joint pain or morning stiffness for more than 30 minutes), negative anti-CCP, no bone erosions on x-rays. Exclusion criteria: People with a known diagnosis of RA by 1987 ACR criteria
Target condition(s)	Rheumatoid arthritis
Index test(s) and reference standard	Ultrasound assessment Scanner: Esaote Mylab 90. All scans performed by rheumatologist trained in musculoskeletal ultrasound and blinded to participant identity and clinical data. 22 joints (in wrists and hands) were scanned. Each scan took at least 15 minutes. Gray-scale synovial hypertrophy and power Doppler synovitis were graded 0-3. Semi quantitative cut-off of GS>1 used for synovial hypertrophy and PD>0 for MCP and PIP joints and PD>1 for wrist joints for synovitis. GS total score and PD total score on 0-66 scale. Presence of tenosynovitis and/or paratendonitis and bone erosions also investigated.

Reference	Ji 2017[17]
	Index test(s) 2010 ACR/EULAR score combined with GS total score 2010 ACR/EULAR score combined with PD total score 2010 ACR/EULAR score combined with synovitis joint count 2010 ACR/EULAR score combined with ≥2 joints with synovitis in the hands Comparator test 2010 ACR/EULAR score Reference standard 1987 ACR criteria after at least 1 year follow-up (median: 15 months). Rheumatologist blinded to US results.
Statistical measures	Index text 2010 ACR/EULAR score combined with GS total score AUC: 0.864 Index text 2010 ACR/EULAR score combined with PD total score AUC: 0.869 Index text 2010 ACR/EULAR score combined with synovitis joint count AUC: 0.872 Index text 2010 ACR/EULAR score combined with ≥2 joints with synovitis in the hands Sensitivity: 0.862 Comparator text 2010 ACR/EULAR score AUC: 0.738
Source of funding	Funded by Capital Health Research and Development of Special and Peking University Clinical Research. Not for profit organisations.
Limitations	Risk of bias: very serious risk of bias due to unclear reporting of index test analysis and selection of participants not indicated to be

Reference	Ji 2017[17]
	consecutive. Indirectness: the study was assessed to be applicable and direct evidence
Comments	

1

Reference	Nakagomi 2013[27]
Study type	Prospective cohort study
Study methodology	Data source: participant data of cohort study Recruitment: Consecutive people with musculoskeletal problems and possible RA diagnosis who were referred to the Department of Allergy and Clinical Immunology at Chiba University Hospital from January 2010 to December 2010.
Number of patients	n = 109
Patient characteristics	Age, mean (SD): 52 (15) Gender (female (%)): 85 (78%) Other relevant characteristics: Unclear treatment at baseline. 41 of 104 (39%) progressed to methotrexate treatment for RA in 1 year. SJC, median (IQR): 1 (0-4) TJC, median (IQR): 1 (0-5) RF positive, n (%): 50 (46%) ACPA positive, n (%): 33 (30%) ESR, mm/h, median (IQR): 18 (7-28) CRP, mg/dl, median (IQR): 1 (0-6) Duration of symptoms ≥6 weeks n (%): 106 (97%) DAS28-CRP, means (SD): 3.08 (1.26) HAQ DI score, median (IQR): 0.5 (0.1-1) Ethnicity: Not detailed. Setting: Department of Allergy and Clinical Immunology at Chiba University Hospital

Reference	Nakagomi 2013[27]
	Country: Japan Inclusion criteria: People with musculoskeletal problems for ≤3 years with possible diagnosis of RA. Possible diagnosis of RA was due to exclusion criteria where musculoskeletal symptoms were explained by other diseases. Exclusion criteria: People whose musculoskeletal symptoms were explained by other diseases or had radiographs of hands and feet that showed erosions typical of RA. People with no clinically swollen joints were not excluded in order to include people with subclinical synovitis.
Target condition(s)	Rheumatoid arthritis
Index test(s) and reference standard	Ultrasound examination Performed on the same day as clinical assessment, radiographs assessed by 1 of 6 rheumatologists trained in musculoskeletal US. The rheumatologist was blinded to the clinical information and laboratory data. Scanner: LOGIQ 7 Pro or LOGIQ E9 or Viamo or Apilo XG or HI VISION Avius or HI VISION Preirus. Power Doppler positivity examination undertaken and graded 0-3 per joint. Synovitis on gray-scale imaging defined on semi quantitative 0-3 scale per joint. Index test 1 2010 ACR/EULAR classification criteria with altered variables to include US. Joint swelling in the classification tree replaced by US detected synovitis. Additionally the joint count in the criteria was determined by the presence of synovitis by US. The scoring required was ≥1 GS ultrasound synovitis. Index test 2 2010 ACR/EULAR classification criteria with altered variables to include US. Joint swelling in the classification tree replaced by US detected synovitis. Additionally the joint count in the criteria was determined by the presence of synovitis by US. The scoring required was ≥2 GS ultrasound synovitis and ≥1 on PD synovitis. Comparator/reference test 2010 ACR/EULAR classification criteria undertaken at the same time point as the index test
Statistical measures	Index text: 2010 ACR/EULAR classification criteria + GS Sensitivity: 82% (67% - 93%) Specificity: 75% (64% - 85%) PPV: 66% NPV: 88%

Reference	Nakagomi 2013[27]
	Index text: 2010 ACR/EULAR classification criteria + GS Sensitivity: 57% (41% - 73%) Specificity: 90% (80% - 96%) PPV: 77% NPV: 78% Index test 1 versus comparator (non US) test Change/reclassification of diagnosis Preliminary diagnosis via comparator test: RA: 40, not-RA: 69 Preliminary diagnosis via index test: RA: 50, not-RA: 59 This was an alteration of preliminary diagnosis 17 people reclassified as having RA after index test. 7 People reclassified as not having RA after index test. Index test 2 versus comparator (non US) test Change/reclassification of diagnosis Preliminary diagnosis via comparator test: RA: 40, not-RA: 69 Preliminary diagnosis via index test: RA: 30, not-RA: 79 This was an alteration of preliminary diagnosis 7 people reclassified as having RA after index test. 17 People reclassified as not having RA after index test
Source of funding	Unclear
Limitations	Risk of bias assessment carried out based on the non-accuracy outcome extracted. Risk of bias: low risk of bias Indirectness: the study was assessed to be applicable and direct evidence
Comments	

1

Reference	Navalho 2013[30]
Study type	Prospective cohort study
Study methodology	Data source: participant data of cohort study Recruitment: consecutive people with untreated clinically apparent synovial swelling at the Hospital da Luz and Hospital de Santa Maria

Reference	Navalho 2013[30]
	in Lisbon. Recruited from April 2009 until February 2012.
Number of patients	n = 45
Patient characteristics	Characteristics occasionally broken down via gold standard into people diagnosed with RA (n=30) after 18 months and those not diagnosed with RA (n=15) after 18 months. Age, median (range): 46 (18-73) Gender (%): 40 (89%) women and 5 (11%) men. Other relevant characteristics: Unclear what treatment was received during follow-up. TJC, median (IQR): RA: 8 (11), non-RA: 3 (3) SJC, median (IQR): RA: 4 (6), non-RA: 1 (2) ESR, mm/h, median (IQR): RA: 28 (24), non-RA: 6 (8) Overall disease activity, VAS, median (IQR): RA: 60 (30), non-RA: 60 (29) DAS28, median (IQR): RA: 4 (6), non-RA: 1 (2) SJC, median (IQR): RA: 5 (2), non-RA: 3 (1) RF positive, n (%): RA: 21 (70%), non-RA: 3 (20%) ACPA positive, n (%): RA: 24 (80%), non-RA: 1 (7%) Ethnicity: Not detailed Setting: Two rheumatology outpatient clinics Country: Portugal Inclusion criteria: Consecutive people with untreated clinically apparent synovial swelling at 4 or more of 68 joint count, including involvement of at least 1 joint of the wrists or hands and with disease duration less than 12 months. Exclusion criteria: pregnant or breastfeeding, inability to give informed consent, current use of glucocorticoids, methotrexate, or other DMARDS, active malignancy, cellulites, occupation or sports related overuse, trauma, contraindications to performing an MRI.

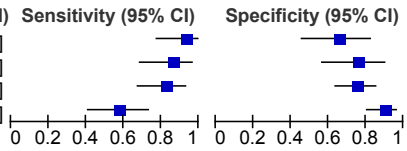
Reference	Navalho 2013[30]
Target condition(s)	Rheumatoid arthritis
Index test(s) and reference standard	US examination GE Logiq 9 scanner with linear array transducer. Undertaken by trained US user blinded to patient's clinical status and MRI results. Evaluation of radioulnar joint, radiocarpal joint, intercarpal and CMC joints, MXP joints, first MCP and first PIP. Also evaluated: tendons, synovial hypertrophy, power Doppler positivity. Synovitis and PD positivity quantified on a 0-3 scale. Index test ACR/EULAR 2010 classification criteria where clinical joint counts were altered to US joint and tendon counts. Comparator test ACR/EULAR 2010 classification criteria Reference standard 1987 ACR criteria at 12 months follow-up.
Statistical measures	Index text ACR/EULAR 2010 classification criteria with US AUC: 0.948 (95% CI: 0.836-0.992) Comparator test (non-US): AUC: 0.909 (95% CI: 0.783-0.975)
Source of funding	Not detailed
Limitations	Risk of bias: low risk of bias Indirectness: the study was assessed to be applicable and direct evidence
Comments	

Appendix E: Coupled sensitivity and specificity forest plots

E.1 Coupled sensitivity and specificity forest plots

Figure 2: ultrasound combined with 1987 ACR criteria

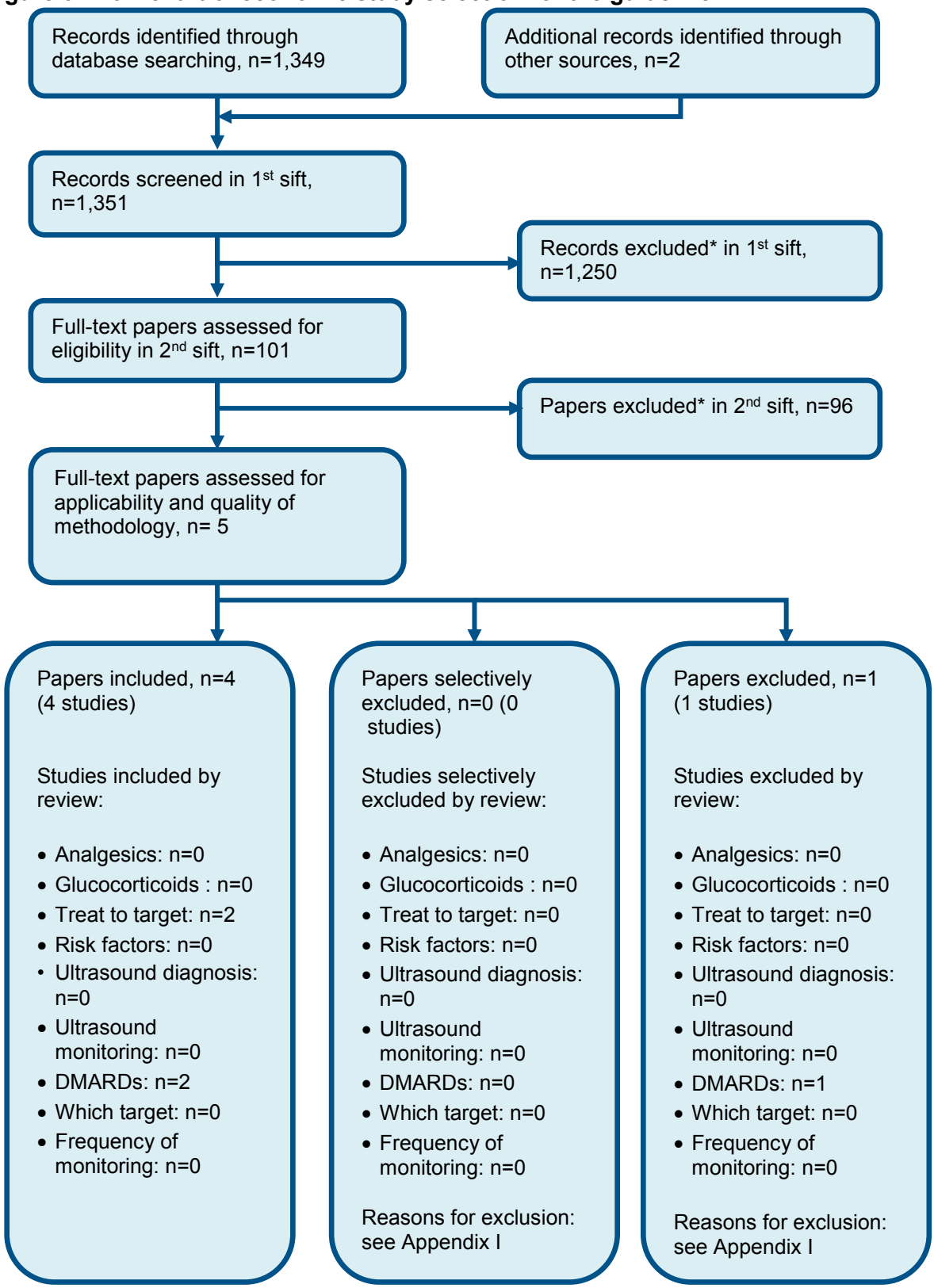
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Filer 2011: GS	27	10	2	19	0.93 [0.77, 0.99]	0.66 [0.46, 0.82]
Filer 2011: PD	25	7	4	22	0.86 [0.68, 0.96]	0.76 [0.56, 0.90]
Nakagomi 2013: GS	33	17	7	52	0.82 [0.67, 0.93]	0.75 [0.64, 0.85]
Nakagomi 2013: GS + PD	23	7	17	62	0.57 [0.41, 0.73]	0.90 [0.80, 0.96]



1

Appendix F:Health economic evidence selection

Figure 3: Flow chart of economic study selection for the guideline



** Non-relevant population, intervention, comparison, design or setting; non-English language*

1

2

3

Appendix G: Health economic evidence tables

None.

1 Appendix H: Excluded studies

H.1 2 Excluded clinical studies

3 **Table 12: Studies excluded from the clinical review**

Reference	Reason for exclusion
Agrawal 2009[1]	No diagnosis of RA data.
Aydin 2017[2]	Unable to obtain paper
Botar-Jid 2010 [4]	Study evaluating ultrasound correlation with blood tests in people with early RA
Broll 2012[5]	No diagnosis of RA data.
Chaiamnuay 2008[6]	Ultrasound assessment not combined with clinical information and laboratory data.
D'Agostino 2016[7]	Not primary research or a systematic review
D'Agostino 2016[8]	Not primary research or a systematic review
El Miedany 2008[9]	Prediction of persistent early inflammatory arthritis
Freeston 2010[11]	Diagnosis of inflammatory arthritis
Ha 2016[12]	Unobtainable
Hirata 2017[13]	Incorrect study design
Hmamouchi 2011[14]	Study evaluating ultrasound detection of flexor tenosynovitis
Horton 2017[15]	Incorrect study design
Hurnakova 2016[16]	Incorrect study design
Kamel 2017[18]	Incorrect study design
Kawashiri 2013[19]	Ultrasound assessment not combined with clinical information and laboratory data.
Komarova 2017[20]	Incorrect study design
Lage-Hansen 2017[21]	Review, not primary research. US diagnostic performance studies checked for inclusion in this review
Lai 2016[22]	Not primary research or a systematic review
Mankia 2016[23]	Not primary research or a systematic review
Mathew 2016[24]	Review, not primary research.US diagnostic performance studies checked for inclusion in this review
Millot 2011[25]	Not a diagnosis of RA study
Minowa 2016[26]	Unobtainable
Naredo 2016[28]	Review, not primary research. RA MSUS diagnostic performance studies checked for inclusion in this review
Ohrndorf 2015[31]	Not primary research
Ozgul 2009[32]	Ultrasound assessment not combined with clinical information and laboratory data.
Plaza 2016[33]	Review, not primary research. US diagnostic performance studies checked for inclusion in this review
Ponikowska 2015[34]	No diagnosis of RA data.

Reference	Reason for exclusion
Pratt 2013[35]	No diagnostic accuracy data on people with and without RA
Rakieh 2015[36]	Diagnosis of inflammatory arthritis
Rezaei 2014[37]	No diagnostic accuracy data
Rizzo 2015[38]	Participants have had rheumatic disease for a mean of over 10 years
Salaffi 2010[39]	No comparative test that differed from the index test only by not utilising ultrasound
Schmidt 2001 [40]	Not primary research or a systematic review
Sizova 2012[41]	Study investigating anti-MCV for diagnosing RA
Sizova 2015[42]	Not primary research
Takase-Minegishi 2017[43]	Systematic review that is not relevant for this evidence review
Tamas 2013 [44]	Ultrasound assessment not combined with clinical information and laboratory data.
Valor 2016[45]	Incorrect study design
van de Stadt 2010[46]	Ultrasound assessment not combined with clinical information and laboratory data.
van der Ven 2017[47]	No relevant outcomes
Zhao 2017[48]	Literature review
Zufferey 2016[49]	Ultrasound assessment not combined with clinical information and laboratory data.

H.2 Excluded health economic studies

Table 13: Studies excluded from the health economic review

Reference	Reason for exclusion
None	

Appendix I: Research recommendations

I.1 Ultrasound in cases of diagnostic uncertainty

Research question: What is the clinical and cost effectiveness of using ultrasound in addition to clinical assessment when there is uncertainty about the diagnosis in adults with suspected rheumatoid arthritis?

Why this is important:

Early diagnosis of RA is essential to reduce the impact of the disease on multiple systems in the body. The course of RA and the initial presentation can be highly variable; most people with RA have definite synovitis on clinical assessment, but sometimes this is not obvious, leading to uncertainty about the diagnosis. Ultrasound is a clinically accessible, non-invasive and relatively inexpensive imaging modality that can detect subclinical synovitis and early erosive disease and may therefore help determine an early diagnosis of RA in those where the diagnosis would otherwise be uncertain. Early diagnosis enables earlier treatment providing an opportunity to improve the longer term outcomes of people with RA. The additional use of ultrasound may also allow healthcare professionals to be more confident about ruling out a diagnosis of RA

Criteria for selecting high-priority research recommendations:

PICO question	Population: Adults with suspected rheumatoid arthritis where the diagnosis is uncertain following clinical assessment. Intervention(s): Ultrasound plus clinical assessment Comparison: Clinical assessment Outcome(s): Disease activity, quality of life, function
Importance to patients or the population	Ultrasound may improve diagnosis in people with suspected RA that is difficult to diagnose. Earlier and more definitive diagnosis would enable earlier treatment hopefully improving quality of life both for people with RA and those in whom the diagnosis can be ruled out and resulting in better long term outcomes. In addition, in some cases people with rheumatoid arthritis are reluctant to accept their diagnosis and commence treatment. Ultrasound may help improve patient outcomes in these circumstances by enabling clinicians to show people objective evidence of their joint inflammation and thereby encourage them to commence appropriate therapy
Relevance to NICE guidance	Current NICE guidance is was unable to make a recommendation on the use of ultrasound in the diagnosis of rheumatoid arthritis. This research may therefore enable the added benefit of ultrasound to be established, informing future guidance in this area.
Relevance to the NHS	If ultrasound was found to be clinically and cost effective in aiding the diagnosing certain subgroups of people with suspected RA, its use may increase in those group of people. Although this may require additional upfront resource increase for example supply of an ultrasound machine) and additional training requirements for rheumatologists or other members of the MDT to implement its use any additional upfront costs may be offset by the downstream savings resulting from a prompt diagnosis and earlier treatment initiation, or early discharge if RA can be ruled out.
National priorities	N/A
Current evidence	The evidence review reported in chapter A identified limited

base	heterogeneous evidence assessing the diagnostic accuracy of ultrasound plus clinical assessment, mostly in people with clinically definite synovitis. The evidence was too limited and of insufficient quality to support any recommendation about the use of ultrasound in diagnosis of RA. No evidence was available in the population in whom the diagnosis is unclear despite prior investigations. It is this population in which ultrasound is thought to potentially add value by identifying subclinical synovitis.
Equality	Ultrasound may be of benefit where synovitis is difficult to assess in case of obesity or extensive deformities.
Study design	Diagnostic randomised controlled trial (RCT) comparing the use of ultrasound in addition to clinical assessment versus clinical assessment alone. People with suspected RA where the diagnosis is uncertain following clinical assessment (for example, people with symptoms of rheumatoid arthritis but without clinically definite synovitis) would be recruited into the trial. Randomised to either usual diagnosis and care, without the use of ultrasound to confirm or refute the diagnosis, or diagnosis aided by ultrasound and then usual care. Management strategies for those diagnosed would be the same in each group (tailored to the individual's needs as per current guidance). All participants (including those discharged or in whom RA was ruled out) would be followed up for 2 years. Outcomes assessed would include disease activity, quality of life and function. This RCT could be cluster randomised to enhance feasibility.
Feasibility	The potential challenges to feasibility include the possible small numbers that would be relevant to recruit, thus this is suggested to be either a cluster randomised RCT, or multicentre to increase recruitment potential. Retention of participants for follow up assessment, particularly in the group not diagnosed with RA may also pose a challenge, therefore this should be considered in designing the trial so that outcome assessment sessions are not too onerous for the participants. Cross-site agreement on US score and technique should also be pre-specified to minimise the risk of bias that this may introduce.
Other comments	Further, ultrasound training is being undertaken by many trainees and other members of the MDT to be used in rheumatology practice, but without the level of evidence to support its clinical and cost effectiveness in diagnosis of RA.
Importance	High: the research is essential to inform future updates of key recommendations in the guideline.

1

2