National Institute for Health and Care Excellence

Consultation

Rheumatoid arthritis (update)

Evidence review I Ultrasound monitoring

NICE guideline CG79

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, where appropriate, their carer or guardian.

Local commissioners and 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 All NICE guidance is subject to regular review and may be updated or withdrawn.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 monitoring

1.1 Review questions:

In adults with rheumatoid arthritis (RA), what is the added value of monitoring disease activity with ultrasound?

In adults with poor prognosis rheumatoid arthritis, what is the added value of monitoring disease activity with ultrasound?

1.2 Introduction

Structural damage can happen quickly in rheumatoid arthritis if inflammation is not efficiently suppressed. The widespread use of strategies that aim for clinical remission or low disease activity has significantly improved the prognosis of rheumatoid arthritis. However, progressive bone erosion and relapses can still occur even in clinical remission.

Ultrasound can detect subclinical synovitis, but it is not known whether the use of ultrasound as part of routine monitoring results in improved patient outcomes.

1.3 PICO table

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

Table 1: PICO characteristics of clinical effectivenss review

Population	Adults with rheumatoid arthritis Studies in adults with poor prognostic factors, patients in remission, and patients with early disease (< 1 year) will be presented separately
Interventions	• Clinical assessment plus ultrasound assssment
Comparison	• Clinical assessment alone
Outcomes	CRITICAL • Disease Activity Score (continuous) at 12 months • Quality of life (continuous) at 12 months • Function (continuous) at 12 months IMPORTANT • Remission (dichotomous) at 12 months • Low disease activity (dichotomous) at 12 months • Relapse (dichotomous) at 12 months • Flare (dichotomous) at 12 months • Pain (continuous) at 12 months • Radiographic progression (continuous) at 12 months • Change in planned management at time of testing (dichotomous) over duration of trial • Withdrawal from trial / adherence to strategy (dichotomous) over duration of trial
Study design	Randomised controlled trials (RCTs)

	Systematic Review / Network Meta-Analysis of RCTs
Table 2: PICO characteristics of prognostic question**Population**	Adults with RA. Studies in adults with poor prognostic factors, patients in remission, and patients with early disease (< 1 year) will be presented separately.
Prognostic factor(s)	Prognostic factors: • Ultrasound findings (for example: synovitis, persistent effusion, tendon injuries, synovitis, Doppler flow, tethering of the tendons, ligament, tenosynovitis, erosions) • Disease activity
Outcome(s)	• Disease Activity Score (continuous) • Quality of life (continuous) • Function (continuous) • Remission (dichotomous) • Low disease activity (dichotomous) • Relapse (dichotomous) • Flare (dichotomous) • Pain (continuous) • Radiographic progression (continuous) • Change in planned management at time of testing (dichotomous) • Withdrawal from trial / adherence to strategy (dichotomous) •
Study design	Prospective cohort studies (prognostic) Systematic reviews of the above

This review sought to investigate clinical assessment plus ultrasound using 2 components. Firstly the review sought out randomised cotrolled trials comparing monitoring with clinical assessment combined with ultrasound versus monitoring via clinical assessment alone. The outcomes would give a comparison of the clinical affectiveness of the monitoring mthods.

The second component assessed the prognostic value of monitoring to predict the outcomes of interest. The prognostic factors of interest encompassed clinical assessment and ultrasound through measurement of disease activity and ultrasound variables. Multivariate analysis was utilised to assess whether single factors involved in the assessment were independently associated with the outcomes of interest. Factors such as tender joint count, swollen joint count, pain, C-reactive protein (CRP), erythrocyte sedimentation rate (ESR) and anti-CCP are measures utilised in disease assessment so they have been reported as factors in the evidence.

1.4 Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual.[ref to be added] Methods specific to this review question are described in the review protocol in appendix A.

Declarations of interest were recorded according to NICE's 2014 conflicts of interest policy.

1.5 Clinical evidence

1.5.1 Included studies

A search was conducted for randomised controlled trials and systematic reviews of these study types in the first instance. However as evidence was limited, prospective prognostic cohort studies were also searched for.

One randomised controlled trial and 5 prognostic studies were included in the review; they are summarised in Table 3 and Table 4 below. Evidence from these studies is summarised in the clinical evidence summaries below (see Table 6, Table 7, Table 8, Table 9, Table 10 and Table 11).

See also the study selection flow chart in appendix C, study evidence tables in appendix D, forest plots in appendix E and GRADE tables in appendix H.

1.5.2 Excluded studies

See the excluded studies list in appendix I.

1.5.3 Summary of clinical studies included in the evidence review

Table 3: Summary of randomised controlled trials included in the evidence review

Study	Intervention and comparison	Population	Outcomes	Comments
Haavardsholm 2016[25]	Monitoring of treat-to-target tight control regime using clinical assessment and ultrasound versus monitoring using clinical assessment alone (usual care). US Group: Target was clinical remission (DAS <1.6 and no swollen joints) and no power Doppler (PD) signal in any of the joints assessed. Control group: Target was clinical remission (DAS < 1.6 and no swollen joints).	Adults aged 18-75 years with early RA n=238	• Disease activity score • Rheumatoid Arthritis Impact of Disease (RAID) score • Quality of life (EQ-5D) • Remission • Pain • Radiological progression • Withdrawal from trial due to "no longer willing" • Withdrawal from trial due to adverse events	Participants in both groups were treated according to the same fixed treatment algorithm, adhering to a treat-to-target strategy with DMARD escalation therapy if target was not met. US results could be used by clinicians to overrule DAS-based target decisions if indicated.

Table 4: Summary of prospective cohort prognostic studies included in the evidence review

Study	Population	Analysis	Prognostic variable(s)	Confounders	Outcome	Limitations
Boyesen 2011[5]; Haavardsholm 2008[26]	Adults with early rheumatoid arthritis (<1 year) n = 84	Stepwise multiple logistic regression	Ultrasound grey-scale (USGS) inflammation, DAS28, presence of anti-cyclic citrullinated peptide antibody (anti-CCP+), van der Heijde/Sharp score	All prognostic variables plus an additional eleven clinical, biomarker and demographic variables	Magnetic resonance imaging (MRI) erosive progression	Low risk of bias
Geng 2016[23]	Adults with rheumatoid arthritis in clinical remission n = 126	Multivariable logistic regression	DAS28-ESR (erythrocyte sedimentation rate), PD>0, PD total score and synovial hypertrophy (SH) total score	Additional variables included in univariate analysis not known	Relapse (DAS28-ESR > 2.6)	Very high risk of bias (outcome measurement – blinding not reported; statistical analysis and reporting – univariate analysis (UVA) analysis not reported and multivariate analysis (MVA) model unclear).
Horton 2016[32]	Adults with rheumatoid arthritis (DAS28-CRP4v ≥2.6) n = 217	Multivariable logistic regression	Tender joint count of 28 joints (TJC28), swollen joint count of 28 joints (SJC28), C-reactive protein (CRP) mg/L, patient visual analogue scale (VAS) global disease assessmen	All prognostic variables plus an additional eleven clinical, biomarker and demographic variables	Remission: both DAS28-CRP4v < 2.6 and DAS44-CRP4v < 1.6	Very high risk of bias (study attrition – 52% patients lost to follow up or excluded; outcome measurement – blinding not reported; statistical analysis and reporting – only included variables in MVA if p <

Study	Population	Analysis	Prognostic variable(s)	Confounders	Outcome	Limitations
			t (components of DAS28-CRP), total GS score on US, total power Doppler activity (PDA) score on US			0.05 in UVA, may have missed important variables).
Saleem 2012[60]	Adults with rheumatoid arthritis in remission (no flares in the last 6 months, stable treatment) n = 93	Multiple binary logistic regression	PD present, remission (DAS28<2.6)	All prognostic variables plus health assessment questionnaire disability index (HAQ-DI)	Disease flare	Very high risk of bias (study attrition- no information on drop-outs; prognostic factor measurement- crude measurement; study confounding - few possible confounders accounted for)
Zavada 2017[76]	Adults with early or established rheumatoid arthritis n = 185	Multiple logistic regression	Previous DAS28-CRP and previous US7 assessment (grey-scale synovitis sum score (GSsynSS), power Doppler synovitis sum score (PDsynSS), grey-scale tenosynovitis sum score (GStenSS), power Doppler tenosynovitis sum score (PDtenSS),	All prognostic variables plus previous HAQ	HAQ score (after measuring prognostic variables)	High risk of bias (study attrition- missing data unclear) Serious indirectness (MVA model looks at associations between different time points rather than baseline to 12 months)

Study	Population	Analysis	Prognostic variable(s)	Confounders	Outcome	Limitations
			erosion score (ES)			

See appendix D for full evidence tables.

1.5.4 Quality assessment of clinical studies included in the evidence review

1.5.4.1 RCTs

Table 5: Clinical evidence summary: clinical assessment and ultrasound versus clinical assessment alone

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Monitoring including ultrasound (95% CI)
Disease Activity Score Change in DAS28. Scale from: 0 to 9.4.	230 (1 study) 12 months	MODERATE[1] due to risk of bias		The mean change in Disease Activity Score in the control groups was -2.4	The mean change in Disease Activity Score in the intervention groups was 0 higher (0.35 lower to 0.35 higher)
Function Change in rheumatoid arthritis impact of disease (RAID) score. Scale from: 0 to 10.	230 (1 study) 12 months	MODERATE[1] due to risk of bias		The mean change in rheumatoid arthritis impact of disease score in the control groups was -2.4	The mean change in rheumatoid arthritis impact of disease score in the intervention groups was 0.2 lower (0.76 lower to 0.36 higher)
Quality of life Change in EQ-5D. Scale from: -0.59 to 1.	230 (1 study) 12 months	MODERATE[1] due to risk of bias		The mean change in quality of life in the control groups was 0.25	The mean change in quality of life in the intervention groups was 0.03 higher (0.04 lower to 0.1 higher)
Remission DAS <1.6	230 (1 study) 12 months	LOW[1,2] due to risk of bias, imprecision	RR 0.89 (0.75 to 1.06)	723 per 1000	80 fewer per 1000 (from 181 fewer to 43 more)
Pain Change in VAS. Scale from: 0 to 100.	230 (1 study) 12 months	MODERATE[1] due to risk of bias		The mean change in pain in the control groups was -29.2	The mean change in pain in the intervention groups was 3.3 lower (10.16 lower to 3.56 higher)
Radiological progression	230 (1 study) 24 months	MODERATE[3]		The median (interquartile range - IQR) change in Sharp score in the control	The median change in Sharp score in the intervention groups was 0.5 lower[4] (median (IQR): 1.0 (0-2.5)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Monitoring including ultrasound (95% CI)
Change in Sharp score. Scale from: 0 to 448.				group was 1.5 (0.5-3.0)	
Withdrawal from trial due to "no longer willing"	211 (1 study) 24 months	VERY LOW[1,2] due to risk of bias, imprecision	RR 0.4 (0.08 to 2)	48 per 1000	29 fewer per 1000 (from 44 fewer to 48 more)
Withdrawal from trial due to adverse events	215 (1 study) 24 months	VERY LOW[1,2] due to risk of bias, imprecision	RR 1.15 (0.36 to 3.64)	48 per 1000	7 more per 1000 (from 30 fewer to 126 more)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
3 Imprecision could not be assessed because non-parametric statistics were reported. The confidence interval is relatively wide.
4 Confidence intervals not calculated due to non-parametric statistics

1.5.4.2 Prognostic studies

Table 6: Clinical evidence summary: MRI erosive progression (dichotomous – at 1 year)

Risk factor for predicting MRI erosive progression	Number of studies	Effect (95% CI)	Imprecision	GRADE Quality
USGS inflammation (<0.5 vs ≥0.5)	1	Adjusted OR: 2.01 (1.14-3.53)	none	HIGH
DAS28	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
Anti-CCP+	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
van der Heijde/Sharp score	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a

n/a: unable to assess as data not reported (factor not independently associated with the outcome following multivariable analysis)

Table 7: Clinical evidence summary: Disease flare (increase in disease activity requiring an initiation, change or increase in therapy based on DAS28; dichotomous – at 1 year)

Risk factor for predicting disease flare	Number of studies	Effect (95% CI)	Imprecision	GRADE Quality
Remission (DAS28<2.6)	1	Adjusted OR: 2.71 (0.73-10.14)	serious	VERY LOW
PD present	1	Adjusted OR: 7.57 (1.75-32.76)	none	LOW

Table 8: Clinical evidence summary: Relapse (dichotomous – at 1 year; DAS28-ESR>2.6 following a period of clinical remission)

Risk factor for predicting relapse	Number of studies	Effect (95% CI)	Imprecision	GRADE Quality
PD>0	1	Adjusted OR: 8.8 (2.7 – 28.4)	none	LOW
PD total score	1	Adjusted OR: 1.4 (0.9 – 2.0)	serious	VERY LOW
SH total score	1	Adjusted OR: 0.7 (0.5 – 1.0)	serious	VERY LOW
DAS28-ESR	1	Data unavailable[1]	not assessed	not assessed

[1] Variable was independently associated with the outcome but data is not presented here as it was incorrectly reported by the authors.

Table 9: Clinical evidence summary: Remission (DAS28-CRP4v <2.6; dichotomous – at 1 year)

Risk factor for predicting remission	Number of studies	Effect (95% CI)	Imprecision	GRADE Quality
CRP mg/l	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
Patient VAS global disease assessment	1	Adjusted OR: 0.98 (0.95 – 1.00)	serious	VERY LOW
SJC28	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
TJC28	1	Adjusted OR: 0.93 (0.85 – 1.02)	serious	VERY LOW
Total GS score on US	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
Total PDA score on US	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a

1 n/a: unable to assess as data not reported (factor not independently associated with the outcome following multivariable analysis)

2 **Table 10: Clinical evidence summary: Remission (DAS44-CRP4v <1.6; dichotomous – at 1 year)**

Risk factor for predicting remission	Number of studies	Effect (95% CI)	Imprecision	GRADE Quality
CRP mg/l	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
Patient VAS global disease assessment	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
SJC28	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
TJC28	1	Adjusted OR: 0.88 (0.79 – 0.98)	none	LOW
Total GS score on US	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a
Total PDA score on US	1	Not independently associated with the outcome following multivariable analysis.	n/a	n/a

3 n/a: unable to assess as data not reported (factor not independently associated with the outcome following multivariable analysis)

4 **Table 11: Clinical evidence summary: Function (HAQ score; continuous –1 year after prognostic variables were measured)**

Risk factor for predicting function (at 12 months)	Number of studies	Effect (95% CI)	Imprecision	GRADE Quality
Previous DAS28-CRP	1	Coefficient: 0.161 (0.113 to 0.208)	none	LOW
Previous GSsynSS	1	Coefficient: -0.004 (-0.019 to 0.010)	serious	VERY LOW
Previous PDsynSS	1	Coefficient: -0.021 (-0.040 to -0.002)	none	LOW
Previous GStenSS	1	Coefficient: 0.000 (-0.085 to 0.085)	serious	VERY LOW
Previous PDtenSS	1	Coefficient: -0.015 (-0.078 to 0.048)	serious	VERY LOW
Previous erosion score	1	Coefficient: 0.012 (-0.022 to 0.046)	serious	VERY LOW

5 See appendix F for full GRADE tables.

1.6 Economic evidence

1.6.1 Included studies

No relevant health economic studies were identified.

1.6.2 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 G.

1.6.3 Unit costs

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

Table 12: 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 13: 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.

1.7 Resource costs

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

1.8 Evidence statements

1.8.1 Clinical evidence statements

Evidence from 1 RCT showed no clinically important difference between monitoring with or without ultrasound in terms of disease activity, quality of life, function, remission, pain or withdrawal due to adverse events. Monitoring with ultrasound showed a small benefit in terms of lower radiological progression and fewer withdrawals from the trial due to willingness to participate; however, there was considerable uncertainty in the direction of these effects, limiting the ability to draw firm conclusions (1 to 2 years, moderate quality, n=230).

Evidence from 5 prognostic studies collectively reporting on 6 different outcomes at 1 year was highly inconsistent. Ultrasound grey scale inflammation was reported by 1 study (n=93) to be independently associated with subsequent MRI erosive progression, but was not found to be independently associated with remission (1 study, n=217) or function (1 study, n=93), and synovial hypertrophy was found to be independently associated with a reduced risk of relapse. Power Doppler (PD) measures were independently associated with disease flare in 1 study (n=126), in 2 studies (n=185 and 217) the association between PD and the outcome (relapse or function) was present for some PD measures but not others, and other studies found no independent association between PD and remission (n=217) or function (n=185) (high to very low quality evidence). No single ultrasound factor was found to consistently predict the outcomes utilised by the studies included in the review.

1.8.2 Health economic evidence statements

- No relevant economic evaluations were identified.

1.9 Recommendations

I1. Do not use ultrasound for routine monitoring of disease activity in adults with RA.

1.9.1 Research recommendations

I.RR1. What is the clinical and cost effectiveness of using ultrasound to monitor disease in adults with RA when clinical examination is inconclusive or inconsistent with other signs of disease activity?

1.10 Rationale and impact

1.10.1 Why the committee made the recommendations

Randomised controlled evidence did not support using ultrasound for routine monitoring of RA. However, in the committee's experience ultrasound can be useful for monitoring when clinical examination is inconclusive or is inconsistent with other signs of disease activity (for example, pain or markers of inflammation). The committee decided to make a research recommendation to inform future guidance about using ultrasound in these situations.

1.10.2 Why we need recommendations on this topic

Structural damage can happen quickly in rheumatoid arthritis if inflammation is not efficiently suppressed. The widespread use of strategies that aim for clinical remission or low disease activity has significantly improved the prognosis of rheumatoid arthritis. However, progressive bone erosion and relapses can still occur even in clinical remission.

Ultrasound can detect subclinical synovitis, but it is not known whether the use of ultrasound as part of routine monitoring results in improved patient outcomes.

1.10.2.1 Impact

Use and availability of ultrasound varies widely across the country and even between healthcare professionals in the same department. Some healthcare professionals use it routinely whereas others use it on a case-by-case basis. The recommendation should reduce the overall use of ultrasound while still allowing its use for selected subgroups.

1.11 The committee's discussion of the evidence

1.11.1 Interpreting the evidence

1.11.1.1 The outcomes that matter most

Ultrasound is used to monitor disease activity in addition to clinical assessment; therefore, the most critical outcome was agreed to be the DAS. Other critical outcomes were agreed as quality of life and function.

The important outcomes were agreed as the number of people in remission, and low disease activity, using DAS thresholds. The committee agreed that data reported in this format is not as informative as continuous DAS data but still gives an indication of symptom relief and disease activity improvement. Other important outcomes were the number of people

experiencing a relapse, flare, a change in planned management at time of testing, or withdrew from or adhered to trial, as well as the level of pain and radiographic progression.

For most outcomes, 12-month data was sought apart from radiographic progression where the longest reported time point was reported. Outcomes of change in planned management at the time of testing and withdrawal from trial data from the duration of the entire trial was of interest.

No data were available for the outcomes of low disease activity and change in planned management at the time of testing.

1.11.1.2 The quality of the evidence

The review included a single randomised controlled trial (RCT). The evidence was moderate quality for all of the critical outcomes (change in DAS, change in Rheumatoid Arthritis Impact of Disease Score (RAID), and change in quality of life (EQ-5D)) as well as the important outcomes of change in radiological progression and change in pain. The evidence was generally at high risk of bias due to the absence of blinding in the study and the subjective nature of the outcomes reported; only radiographic progression was at low risk of bias.

Further evidence was available from 5 prognostic studies, but this was generally of low to very low quality and could not be pooled as each study reported different outcomes. The majority of evidence was considered to be at very serious risk of bias. The committee discussed 1 study that reported that 52% of participants were lost to follow up or excluded from the analysis due to their outcome data being incomplete without explanation for the missing data. Many of the other studies also had high levels of participants lost to follow up, or failed to report missing data at all, which reduced the committee's certainty in the results.

Another limitation was that many studies failed to report key aspects of their statistical methods. For many of the outcomes, there were small numbers of participants and low numbers of events, resulting in wide confidence intervals, meaning there was considerable uncertainty as to whether the factor was associated with better or poorer outcomes. The impact of these limitations of the evidence was that the committee agreed they could not place much weight on the data from the prognostic studies. No studies were found that looked at people with poor prognostic factors alone or sperately presented their data. .

1.11.1.3 Benefits and harms

The data from the RCT provided moderate quality evidence that ultrasound made no clinical difference for any of the critical outcomes (disease activity score, RAID score and quality of life) when compared to monitoring without the use of ultrasound. Most of the important outcomes (DAS remission, pain and withdrawal due to adverse events) also failed to show a clinical difference between the groups. The outcomes which did show limited benefit with the use of ultrasound were associated with considerable uncertainty and were inconsistent with the majority of the data that informed the review. The committee therefore placed little weight on the evidence for these outcomes in their deliberations.

The committee discussed the findings from the prognostic studies. The committee could not reconcile the highly inconsistent findings between the outcomes and even between different ultrasound measures for the same outcome. For example, given the association seen between grey scale inflammation and MRI erosive progression, the committee were surprised to see no association between the same factor and the outcomes of remission and function.

The committee agreed that given the limited data (1 small study for each outcome), the low to very low evidence quality, and the inconsistent results across the outcomes, little weight should be placed on the prognostic data in determining the value of ultrasound in monitoring rheumatoid arthritis. The committee agreed that RCT data was the best way to establish the

true added value of monitoring rheumatoid arthritis using ultrasound. In the presence of the clear RCT findings, the committee placed little weight on the (inconsistent, inconclusive) results of the non-randomised prognostic studies.

Overall, the committee agreed that there was no evidence that ultrasound added value over monitoring disease activity clinically for the majority of rheumatoid arthritis patients. As a result, the committee agreed that for most people with rheumatoid arthritis, clinical assessment performs well and there was no reason to recommend the use of ultrasound as it does not provide additional information that would change management or outcomes.

The committee decided that the evidence did not support the routine use of ultrasound in monitoring the majority of people with rheumatoid arthritis and made a recommendation accordingly.

The committee agreed, however, that there may be a proportion of people who might benefit from ultrasound assessment, but these populations were not defined in any of the included studies. These might be people with rheumatoid arthritis where decisions have to be made about escalating treatment, and in whom:

- the clinician perceives a difference between the clinical examination and the disease activity score (for example, where no clinical synovitis is apparent but other markers of disease activity such as inflammatory markers or pain are high); or
- clinical examination is unreliable or uncertain (for example, the evaluation of synovial swelling is affected by other co-existent factors such as obesity or oedema).

The committee noted that the included studies did not reflect any of the potential population's subgroups of interest. Therefore, the committee agreed to make a research recommendation to determine whether ultrasound assessment would add value to standard clinical monitoring in specific subgroups with rheumatoid arthritis.

1.11.2 Cost effectiveness and resource use

No health economic studies were identified. The unit cost of ultrasound (£55 per ultrasound) was presented to the committee to aid the consideration of cost-effectiveness. The committee noted that, in some areas ultrasound monitoring is carried out in the rheumatology department and in other areas it is referred to the radiology department. The unit cost presented to the committee was deemed to reflect the cost of ultrasound within a radiology department appropriately.

The committee discussed the potential economic benefits of ultrasound monitoring in a subset of people. The committee noted that in this subset of people in whom ultrasound monitoring could help to identify remission, treatment could be tapered off, and there could be a reduction in the use of DMARDs. This could potentially offset the cost of monitoring or even save costs to the NHS. The committee conceded that no evidence is currently available to support this.

The clinical evidence did not support the routine use of ultrasound in monitoring the majority of people with rheumatoid arthritis, and the committee made a recommendation accordingly. As routine use of ultrasound for monitoring is used in some rheumatology departments in the NHS, it is anticipated that this recommendation should reduce the overall use of ultrasound and therefore moderately reduce costs to the NHS.

1.11.3 Other factors the committee took into account

The committee discussed that results seen from ultrasound can be meaningful for people with rheumatoid arthritis. The patient representatives on the committee explained that ultrasound enables patients to visualise their disease activity, instead of only been given a score (for example, DAS). This may be reflected by data from the RCT which showed that more people continued to be willing to participate in the ultrasound treatment arm compared

to the clinical monitoring only arm (though the committee again noted the imprecision of this effect estimate). In some circumstances, this visualisation of disease activity may be important and may improve patient outcomes, by encouraging medication adherence and facilitating agreement to treatment escalation where necessary. However, the committee agreed that in the presence of evidence that including ultrasound as part of regular monitoring does not improve clinical outcomes, its routine use could not be justified. Further research should help to clarify the circumstances where ultrasound assessment may be clinically and cost effective in rheumatoid arthritis.

1

References

1. 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; 16:16

2. Backhaus TM, Ohrndorf S, Kellner H, Strunk J, Hartung W, Sattler H et al. The US7 score is sensitive to change in a large cohort of patients with rheumatoid arthritis over 12 months of therapy. Annals of the Rheumatic Diseases. 2013; 72(7):1163-1169

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

4. Bellis E, Scire CA, Carrara G, Adinolfi A, Batticciotto A, Bortoluzzi A et al. Ultrasound-detected tenosynovitis independently associates with patient-reported flare in patients with rheumatoid arthritis in clinical remission: results from the observational study STARTER of the Italian Society for Rheumatology. Rheumatology. 2016; 55(10):1826-1836

5. Boyesen P, Haavardsholm EA, Van Der Heijde D, Ostergaard M, Hammer HB, Sesseng S et al. Prediction of MRI erosive progression: A comparison of modern imaging modalities in early rheumatoid arthritis patients. Annals of the Rheumatic Diseases. 2011; 70(1):176-179

6. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. Arthritis & Rheumatism. 2008; 58(10):2958-2967

7. Bugatti S, Bogliolo L, Vitolo B, Manzo A, Montecucco C, Caporali R. Anti-citrullinated protein antibodies and high levels of rheumatoid factor are associated with systemic bone loss in patients with early untreated rheumatoid arthritis. Arthritis Research & Therapy. 2016; 18:226

8. Bugatti S, Manzo A, Benaglio F, Klersy C, Vitolo B, Todoerti M et al. Serum levels of CXCL13 are associated with ultrasonographic synovitis and predict power Doppler persistence in early rheumatoid arthritis treated with non-biological disease-modifying anti-rheumatic drugs. Arthritis Research & Therapy. 2012; 14:R34

9. Cavet G, Shen Y, Abraham S, Chernoff D, Centola M, Taylor P. Predicting radiographic progression in rheumatoid arthritis with ultrasound and biomarkers. Arthritis & Rheumatism. 2009; 60(Suppl 10):1464

10. Chen YC, Su FM, Hsu SW, Chen JF, Cheng TT, Lai HM et al. Predictor of hand radiological progression in patients with rheumatoid arthritis receiving TNF antagonist therapy by change in grayscale synovitis-a preliminary study. Journal of Clinical Rheumatology. 2017; 23(2):73-76

11. Cheung P, Mari K, Devauchelle V, Bentin J, Jousse-Joulin S, D'Agostino MA et al. Are tender joints better than synovitis to predict structural damage in rheumatoid arthritis? Arthritis & Rheumatology. 2014; 66(Suppl 10):S602

12. Cheung PP, Mari K, Devauchelle-Pensec V, Jousse-Joulin S, D'Agostino MA, Chales G et al. Predictive value of tender joints compared to synovitis for structural damage in rheumatoid arthritis. RMD Open. 2016; 2(1):e000205

13. Christensen AW, Rifbjerg-Madsen S, Christensen R, Dreyer L, Boesen M, Ellegaard K et al. Ultrasound Doppler but not temporal summation of pain predicts DAS28

response in rheumatoid arthritis: a prospective cohort study. Rheumatology. 2016; 55(6):1091-1098

14. D'Agostino MA, Boers M, Wakefield RJ, Berner Hammer H, Vittecoq O, Filippou G et al. Exploring a new ultrasound score as a clinical predictive tool in patients with rheumatoid arthritis starting abatacept: results from the APPRAISE study. RMD Open. 2016; 2(1):e000237

15. D'Agostino MA, Boers M, Wakefield RJ, Emery P, Conaghan PG. Is it time to revisit the role of ultrasound in rheumatoid arthritis management? Annals of the Rheumatic Diseases. 2017; 76(1):7-8

16. Dale J, Purves D, McConnachie A, McInnes I, Porter D. Tightening up? Impact of musculoskeletal ultrasound disease activity assessment on early rheumatoid arthritis patients treated using a treat to target strategy. Arthritis Care & Research. 2014; 66(1):19-26

17. Dale J, Stirling A. Targeting ultrasound remission in early rheumatoid arthritis: results of the TaSER study, a randomised clinical trial. Annals of the Rheumatic Diseases. 2016; 75(6):1043-1050

18. Dougados M, Devauchelle-Pensec V, Ferlet JF, Jousse-Joulin S, D'Agostino MA, Backhaus M et al. Both clinical and ultrasonographic evaluation of synovitis are relevant to predict subsequent radiological deterioration in rheumatoid arthritis. Annals of the Rheumatic Diseases. 2012; 71(Suppl 3):610

19. Dougados M, Devauchelle-Pensec V, Ferlet JF, Jousse-Joulin S, D'Agostino MA, Backhaus M et al. The ability of synovitis to predict structural damage in rheumatoid arthritis: a comparative study between clinical examination and ultrasound. Annals of the Rheumatic Diseases. 2013; 72(5):665-671

20. El Miedany Y, El Gaafary M, Youssef S, Ahmed I, Bahlas S, Hegazi M et al. Optimizing therapy in inflammatory arthritis: prediction of relapse after tapering or stopping treatment for rheumatoid arthritis patients achieving clinical and radiological remission. Clinical Rheumatology. 2016; 35(12):2915-2923

21. Fukae J, Tanimura K, Isobe M, Kitano A, Henmi M, Nakai M et al. Active synovitis in the presence of osteitis predicts residual synovitis in patients with rheumatoid arthritis with a clinical response to treatment. International Journal of Rheumatic Diseases. 2016; Epublication

22. Gartner M, Alasti F, Supp G, Mandl P, Smolen JS, Aletaha D. Persistence of subclinical sonographic joint activity in rheumatoid arthritis in sustained clinical remission. Annals of the Rheumatic Diseases. 2015; 74(11):2050-2053

23. Geng Y, Han J, Deng X, Zhang Z. Deep clinical remission: an optimised target in the management of rheumatoid arthritis? Experience from an ultrasonography study. Clinical and Experimental Rheumatology. 2016; 34(4):581-586

24. Haavardsholm EA, Aga AB, Olsen IC, Hammer HB, Uhlig T, Fremstad H et al. Aiming for remission in rheumatoid arthritis - The ARCTIC Trial. Arthritis & Rheumatology. 2016; 68(Suppl 10):10L

25. Haavardsholm EA, Aga AB, Olsen IC, Lillegraven S, Hammer HB, Uhlig T et al. Ultrasound in management of rheumatoid arthritis: ARCTIC randomised controlled strategy trial. BMJ. 2016; 354:i4205

26. Haavardsholm EA, Boyesen P, Ostergaard M, Schildvold A, Kvien TK. Magnetic resonance imaging findings in 84 patients with early rheumatoid arthritis: bone

marrow oedema predicts erosive progression. Annals of the Rheumatic Diseases. 2008; 67(6):794-800

27. Hammer HB, Kvien TK, Terslev L. Ultrasound of the hand is sufficient to detect subclinical inflammation in rheumatoid arthritis remission: a post hoc longitudinal study. Arthritis Research & Therapy. 2017; 19(1):221

28. Han J, Geng Y, Deng X, Zhang Z. Subclinical synovitis assessed by ultrasound predicts flare and progressive bone erosion in rheumatoid arthritis patients with clinical remission: A systematic review and metaanalysis. Journal of Rheumatology. 2016; 43(11):2010-2018

29. Harman H, Tekeoglu I, Kaban N, Harman S. Factors influencing ultrasonographic remission in patients with rheumatoid arthritis. Rheumatology International. 2015; 35(3):485-491

30. Harman H, Tekeoglu I, Takci S, Kamanli A, Nas K, Harman S. Improvement of large-joint ultrasonographic synovitis is delayed in patients with newly diagnosed rheumatoid arthritis: results of a 12-month clinical and ultrasonographic follow-up study of a local cohort. Clinical Rheumatology. 2015; 34(8):1367-1374

31. 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

32. Horton SC, Tan AL, Freeston JE, Wakefield RJ, Buch MH, Emery P. Discordance between the predictors of clinical and imaging remission in patients with early rheumatoid arthritis in clinical practice: implications for the use of ultrasound within a treatment-to-target strategy. Rheumatology. 2016; 55(7):1177-1187

33. 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

34. 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

35. Hurnakova J, Zavada J, Hanova P, Hulejova H, Klein M, Mann H et al. Serum calprotectin (S100A8/9): an independent predictor of ultrasound synovitis in patients with rheumatoid arthritis. Arthritis Research & Therapy. 2015; 17:252

36. Iagnocco A, Finucci A, Ceccarelli F, Perricone C, Iorgoveanu V, Valesini G. Power Doppler ultrasound monitoring of response to anti-tumour necrosis factor alpha treatment in patients with rheumatoid arthritis. Rheumatology. 2015; 54(10):1890-1896

37. Ikeda K, Brown A, Conaghan P, Karim Z, Quinn M, Peterfy C. An explanation of structural deterioration in RA patients in clinical remission on DMARDS: The significance of sub-clinical inflammation detectable with imaging. Annals of the Rheumatic Diseases. 2007; 66(Suppl. 2):95

38. Ikeda K, Nakagomi D, Sanayama Y, Yamagata M, Okubo A. Time-integrated synovitis activity assessed by power doppler ultrasound significantly correlates with radiographic progression in rheumatoid arthritis patients treated with methotrexate

alone but not in those treated with TNF antagonists. Arthritis & Rheumatism. 2012; 64(Suppl 10):S50

39. Ikeda K, Nakagomi D, Sanayama Y, Yamagata M, Okubo A, Iwamoto T et al. Correlation of radiographic progression with the cumulative activity of synovitis estimated by power Doppler ultrasound in rheumatoid arthritis: difference between patients treated with methotrexate and those treated with biological agents. Journal of Rheumatology. 2013; 40(12):1967-1976

40. Ivanac G, Morovic-Vergles J, Brkljacic B. Gray-scale and color duplex Doppler ultrasound of hand joints in the evaluation of disease activity and treatment in rheumatoid arthritis. Croatian Medical Journal. 2015; 56(3):280-289

41. Janta I, Valor L, De la Torre I, Martinez-Estupinan L, Nieto JC, Ovalles-Bonilla JG et al. Ultrasound-detected activity in rheumatoid arthritis on methotrexate therapy: Which joints and tendons should be assessed to predict unstable remission? Rheumatology International. 2016; 36(3):387-396

42. Jeka S, Dura M, Zuchowski P, Zwierko B, Wojciechowski R. The role of ultrasonography in monitoring long-standing rheumatoid arthritis: A pilot study. Reumatologia. 2017; 55(4):177-182

43. Jindal S, Kaushik R, Raghuvanshi S, Kaushik RM, Kakkar R. Gray scale and power Doppler ultrasonographic findings in the assessment of disease activity and their correlation with disease activity parameters in rheumatoid arthritis. Current Rheumatology Reviews. 2017; Epublication

44. Kakavouli GG, Siagkri C, Zivanovic S. Ultrasound imaging of the joint damage in the patient with long-standing rheumatoid arthritis. Hippokratia. 2015; 19(3):281

45. 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

46. Kawashiri SY, Nishino A, Shimizu T, Umeda M, Fukui S, Nakashima Y et al. Ultrasound disease activity of bilateral wrist and finger joints at three months reflects the clinical response at six months of patients with rheumatoid arthritis treated with biologic disease-modifying anti-rheumatic drugs. Modern Rheumatology. 2017; 27(2):252-256

47. Kirino Y, Hama M, Takase-Minegishi K, Kunishita Y, Kishimoto D, Yoshimi R et al. Predicting joint destruction in rheumatoid arthritis with power Doppler, anti-citrullinated peptide antibody, and joint swelling. Modern Rheumatology. 2015; 25(6):842-848

48. 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

49. Luz KR, Pinheiro MM, Petterle GS, Dos Santos MF, Fernandes AR, Natour J et al. A new musculoskeletal ultrasound scoring system (US10) of the hands and wrist joints for evaluation of early rheumatoid arthritis patients. Revista Brasileira de Reumatologia. 2016; 56(5):421-431

50. Manzo A, Bugatti S, Benaglio F, Vitolo B, Klersy C, Montecucco C et al. Serum levels of CXCL13 are associated with ultrasonographic synovitis and predict power doppler persistence in early rheumatoid arthritis treated with non-biological diseasemodifying antirheumatic drugs. Annals of the Rheumatic Diseases. 2012; 71(Suppl 1):A12-A13

51. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P et al. Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. Arthritis & Rheumatism. 2007; 57(1):116-124

52. Naredo E, Moller I, Cruz A, Carmona L, Garrido J. Power Doppler ultrasonographic monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. Arthritis & Rheumatism. 2008; 58(8):2248-2256

53. 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

54. Osipyants R, Karateev D, Panasyuk E, Smirnov A, Lukina G, Glukhova S et al. Associations between functional status and ultrasound-detected synovitis and joint damage in rheumatoid arthritis during tocilizumab treatment. Annals of the Rheumatic Diseases. 2013; 72(Suppl 3):A755-756

55. Osipyants R, Karateev D, Panasyuk E, Smirnov A, Lukina G, Glukhova S et al. Imaging rather than clinical inflammation is associated with radiographic progression in tocilizumab-treated rheumatoid arthritis patients. Annals of the Rheumatic Diseases. 2013; 72(Suppl 3):A755

56. Ramírez García J, Ruíz-Esquide V, Celis R, Cuervo A, Cabrera S, Inciarte-Mundo J et al. AB0254 Sonographic and clinical characterization of a prospective cohort of patients with rheumatoid arthritis in clinical remission. Preliminary results. Annals of the Rheumatic Diseases. 2014; 73(Suppl 2):888

57. Rees JD, Pilcher J, Heron C, Kiely PD. A comparison of clinical vs ultrasound determined synovitis in rheumatoid arthritis utilizing gray-scale, power Doppler and the intravenous microbubble contrast agent 'Sono-Vue'. Rheumatology. 2007; 46(3):454-459

58. Reynolds PPM, Heron C, Pilcher J, Kiely PDW. Prediction of erosion progression using ultrasound in established rheumatoid arthritis: A 2-year follow-up study. Skeletal Radiology. 2009; 38(5):473-478

59. Rosa J, Ruta S, Saucedo C, Navarta DA, Catoggio LJ, Garcia-Monaco R et al. Does a simplified 6-joint ultrasound index correlate well enough with the 28-joint disease activity score to be used in clinical practice? Journal of Clinical Rheumatology. 2016; 22(4):179-183

60. Saleem B, Brown AK, Quinn M, Karim Z, Hensor EM, Conaghan P et al. Can flare be predicted in DMARD treated RA patients in remission, and is it important? A cohort study. Annals of the Rheumatic Diseases. 2012; 71(8):1316-1321

61. Scire CA, Montecucco C, Codullo V, Epis O, Todoerti M, Caporali R. Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. Rheumatology. 2009; 48(9):1092-1097

62. Sreerangaiah D, Grayer M, Fisher BA, Ho M, Abraham S, Taylor PC. Quantitative power Doppler ultrasound measures of peripheral joint synovitis in poor prognosis early rheumatoid arthritis predict radiographic progression. Rheumatology. 2016; 55(1):89-93

63. 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

64. Tan YK, Allen JC, Jr., Lye WK, Conaghan PG, D'Agostino MA, Chew LC et al. Novel ultrasound joint selection methods using a reduced joint number demonstrate inflammatory improvement when compared to existing methods and disease activity score at 28 joints. Journal of Rheumatology. 2016; 43(1):34-37

65. Taylor PC, Steuer A, Gruber J, Cosgrove DO, Blomley MJ, Marsters PA et al. Comparison of ultrasonographic assessment of synovitis and joint vascularity with radiographic evaluation in a randomized, placebo-controlled study of infliximab therapy in early rheumatoid arthritis. Arthritis & Rheumatism. 2004; 50(4):1107-1116

66. Tokai N, Ogasawara M, Gorai M, Matsuki Y, Yamada Y, Murayama G et al. Predictive value of bone destruction and duration of clinical remission for subclinical synovitis in rheumatoid arthritis patients. Modern Rheumatology. 2015; 25(4):540-545

67. Toyota Y, Tamura M, Kirino Y, Sugiyama Y, Tsuchida N, Kunishita Y et al. Musculoskeletal ultrasonography delineates ankle symptoms in rheumatoid arthritis. Modern Rheumatology. 2017; 27(3):425-429

68. 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

69. 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

70. Vlad V, Berghea F, Micu M, Varzaru L, Bojinca M, Milicescu M et al. Tenosynovitis US scoring systems follow synovitis and clinical scoring systems in RA and are responsive to change after biologic therapy. Medical Ultrasonography. 2015; 17(3):352-360

71. Vreju FA, Filippucci E, Gutierrez M, Di Geso L, Ciapetti A, Ciurea ME et al. Subclinical ultrasound synovitis in a particular joint is associated with ultrasound evidence of bone erosions in that same joint in rheumatoid patients in clinical remission. Clinical and Experimental Rheumatology. 2016; 34(4):673-678

72. Wakefield RJ, Freeston JE, Hensor EM, Bryer D, Quinn MA, Emery P. Delay in imaging versus clinical response: a rationale for prolonged treatment with anti-tumor necrosis factor medication in early rheumatoid arthritis. Arthritis & Rheumatism. 2007; 57(8):1564-1567

73. Yamada Y, Ogasawara M, Gorai M, Matsuki Y, Murayama G, Sugisaki N et al. The synovial grade corresponding to clinically involved joints and a feasible ultrasound-adjusted simple disease activity index for monitoring rheumatoid arthritis. Modern Rheumatology. 2016; 26(6):844-849

74. Yoshimi R, Hama M, Minegishi K, Kishimoto D, Watanabe T, Kamiyama R et al. Ultrasonography predicts achievement of Boolean remission after DAS28-based clinical remission of rheumatoid arthritis. Modern Rheumatology. 2014; 24(4):590-598

75. Yoshimi R, Hama M, Takase K, Ihata A, Kishimoto D, Terauchi K et al. Ultrasonography is a potent tool for the prediction of progressive joint destruction during clinical remission of rheumatoid arthritis. Modern Rheumatology. 2013; 23(3):456-465

76. Zavada J, Hanova P, Hurnakova J, Szczukova L, Uher M, Forejtova S et al. The relationship between synovitis quantified by an ultrasound 7-joint inflammation score and physical disability in rheumatoid arthritis - a cohort study. Arthritis Research & Therapy. 2017; 19:5

77. 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

1 Appendices

2 Appendix A: Review protocols

3 **Table 14: Review protocol: Use of ultrasound monitoring in rheumatoid arthritis**

ID	Field	Content
I	Review question	In adults with rheumatoid arthritis, what is the added value of monitoring disease activity with ultrasound? In adults with poor prognosis rheumatoid arthritis, what is the added value of monitoring disease activity with ultrasound?
II	Type of review question	Combined prognostic and clinical effectiveness (intervention) 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	The aim of this review is to determine the clinical and cost effectiveness of using ultrasound in the monitoring of rheumatoid arthritis and the utility of ultrasound findings in predicting worsening disease outcomes.
IV	Eligibility criteria – population / disease / condition / issue / domain	Adults with rheumatoid arthritis according to validated classification criteria. Studies in the following groups of patients will be analysed and reported separately: • poor prognosis rheumatoid arthritis • in remission • early disease (<1 year)
V	Eligibility criteria – intervention(s) / exposure(s) / prognostic factor(s)	Ultrasound assessment and clinical assessment (any joints). Factor of interest: Ultrasound findings (e.g. synovitis, persistent effusion, tendon injuries, synovitis, Doppler flow, tethering of the tendons, ligament, tenosynovitis, erosions), high or increased disease activity Ultrasound assessment should be performed by an appropriately trained healthcare professional.
VI	Eligibility criteria – comparator(s) / control or reference (gold) standard	Clinical assessment only (e.g. using a validated composite disease activity measure)
VII	Outcomes and prioritisation	CRITICAL • Disease Activity Score (continuous) at 12 months • Quality of life (for example, EQ5D, SF-36, RA Quality of Life instrument) (continuous) at 12 months • Function (for example, Health Assessment Questionnaire, activities of daily living) (continuous) at 12 months IMPORTANT • Remission (dichotomous) at 12 months • Low disease activity (dichotomous) at 12 months • Relapse (dichotomous) at 12 months • Flare (dichotomous) at 12 months • Pain (for example, visual analogue scale) (continuous) at 12 months • Radiographic progression (continuous) at 12 months

ID	Field	Content
		• Change in planned management at time of testing (dichotomous) over duration of trial • Withdrawal from trial / adherence to strategy (dichotomous) over duration of trial For outcomes other than radiological progression, data must be reported least 6 months after use of ultrasound has commenced. If multiple time points, take closest time point to 12 months. For radiological progression, data must be at least 12 months. If multiple time points, take the longest time point. Stage 1 of the review will analyse RCT data to establish whether monitoring using ultrasound has an impact on each of the above outcomes. Stage 2 of the review will analyse prognostic data on the relationship between ultrasound-detected changes and each of the above outcomes. The follow outcomes will be extracted: • Change in disease activity (as defined by the study)
VIII	Eligibility criteria – study design	RCTs Prospective cohort studies (prognostic only) Systematic reviews of the above
IX	Other inclusion / exclusion criteria	Studies in mixed inflammatory arthritis populations will be excluded, unless the results are presented separately for RA patients. Studies in patients with RA as well as another rheumatic disease (e.g. lupus) will be excluded. Studies reporting association data in Stage 2 of the review (e.g. odds ratios) must adjust for key confounders (for radiological progression: anti-CPP positivity and baseline erosions; for disease activity-related outcomes: a validated disease activity score or its key components (swollen joints and inflammatory markers)).
X	Proposed sensitivity / subgroup analysis, or meta-regression	Stratification – groups that will be considered separately if data are available: • Small joints (hands, wrists, feet) versus other joints versus mixed small/other joints
XI	Selection process – duplicate screening / selection / analysis	A sample of at least 10% of the abstract lists were double-sifted by a senior research fellow and discrepancies rectified, with committee input where consensus could not be reached, for more information please see the separate Methods report for this guideline.
XII	Data management (software)	• Pairwise meta-analyses will be performed using Cochrane Review Manager (RevMan5). • GRADEpro will be used to assess the quality of evidence for each outcome. • Endnote will be used for bibliography, citations, sifting and reference

ID	Field	Content
		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 amendment to previous protocol	For details, please see section 4.5 of Developing NICE guidelines: the 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	Standard study checklists will be used to appraise individual studies critically. For details please see section 6.2 of Developing NICE guidelines: the manual QUIPS tool will be used for the evaluation of risk of bias for prognostic studies. The risk of bias across all available evidence will be evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/ A modified GRADE approach will be used for prognostic studies.
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	For details, please see section 6.2 of Developing NICE guidelines: the manual.

ID	Field	Content
	bias, selective reporting bias	
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.
XXIX	Roles of sponsor	NICE funds NGC to develop guidelines for those working in the NHS, public health and social care in England.
XXX	PROSPERO registration number	Not registered

1 **Table 15: 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).53

Review question	All questions – health economic evidence
	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: 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

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 16: 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 17: 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)

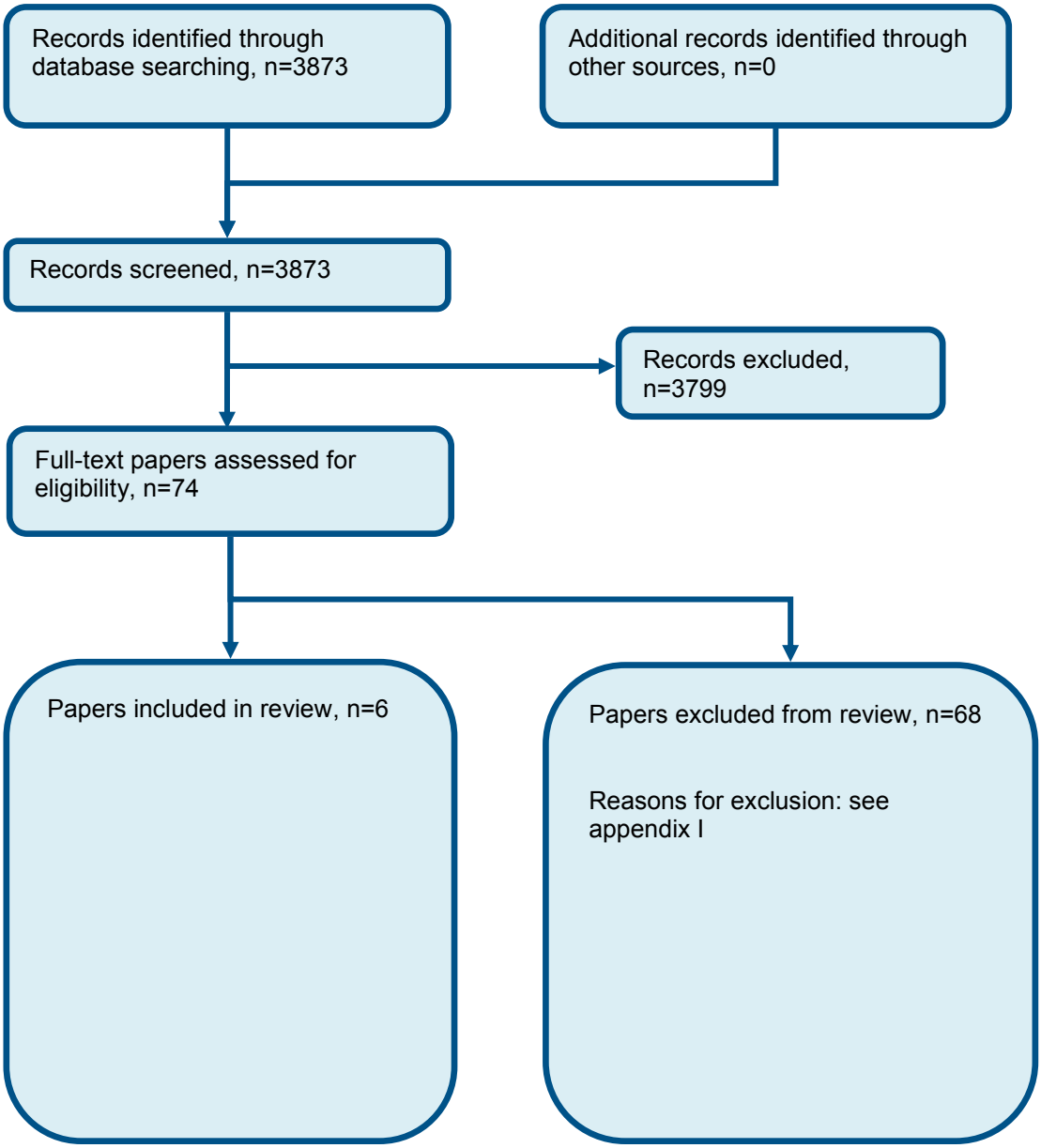
1 **NHS EED and HTA (CRD) search terms**

#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 monitoring



Appendix D: Clinical evidence tables

D.1 Randomised controlled trials

Study (subsidiary papers)	ARCTIC trial trial: Haavardsholm 2016[25] (Haavardsholm 2016[24])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=238)
Countries and setting	Conducted in Norway; Setting: 11 centres in Norway (4 rheumatology departments in university hospitals, 6 regional/community hospitals, and 1 private practice).
Line of therapy	Adjunctive to current care
Duration of study	Intervention + follow up: 24 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: 2010 EULAR criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients had to meet all of the following criteria: 1. Male or non-pregnant, non-nursing female. 2. >18 years of age and <75 years of age. 3. Classified as having RA (according to 2010 ACR/EULAR criteria). 4. Disease duration less than 2 years (defined as time from 1st joint swelling). 5. The treating rheumatologist decided the patient required DMARD-treatment. 6. No prior DMARD use. 7. Patients able and willing to give written informed consent and comply with the requirements of the study protocol.
Exclusion criteria	1. Abnormal renal function (serum creatinine >142 µmol/L in female and >168 µmol/L in male, or GFR <40 mL/min/1.73 m2). 2. Abnormal liver function (ASAT/ALAT > 3* normal), active or recent hepatitis, cirrhosis. 3. Major co-morbidities like severe malignancies, severe diabetic mellitus, severe infections, uncontrollable hypertension, severe cardiovascular disease (NYHA class 3-4) and/or severe respiratory diseases. 4. Leukopenia and/or thrombocytopenia. 5. Inadequate birth control conception, pregnancy, and/or breastfeeding. 6. Indications of active tuberculosis.

	7. Psychiatric or mental disorders, alcohol abuse or other abuse of substances, language barriers or other factors which makes adherence to the study protocol impossible.
Recruitment/selection of patients	The site investigators enrolled patients. Open label, parallel group clinical strategy study. Investigators and patients were aware of the allocated treatment group.
Age, gender and ethnicity	Age - Mean (SD): Intervention group: 50.6 (13.3); Usual care: 52.3 (14.1). Gender (M:F): 2/3. Ethnicity: NR
Further population details	
Extra comments	.
Indirectness of population	No indirectness
Interventions	(n=122) Intervention 1: Monitoring using clinical assessment and ultrasound - Ultrasound monitoring - joints mixed. Assessed at 0, 1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 20, and 24 months. People in the ultrasound strategy arm were assessed by ultrasound at every visit, according to a scoring system of 32 joints with high intra-rater and inter-rater reliability. These 32 joints were scored semi-quantitatively as 0-3 for both grey scale and power Doppler: metacarpophalangeal joints 1-5, radiocarpal joint, intercarpal joint, distal radioulnar joint, elbow, knee, talocrural joint, and metatarsophalangeal joints 1-5 bilaterally, giving ranges from 0 to 192 for total ultrasound score and from 0 to 96 for grey scale and power Doppler ultrasound scores. In the ultrasound strategy arm, the sonographer was also the treating physician, and patients were informed of the ultrasound results. The treatment target in the ultrasound tight control strategy was clinical remission (defined as Disease Activity Score <1.6 and no swollen joints) and ultrasound imaging remission (defined as no power Doppler signal in any of the joints assessed by ultrasound). The ultrasound standardised score included assessments of the following 32 joints with both grey-scale and power Doppler (semi-quantitative score of 0-3 for all joints, with a reference atlas showing the different possible grades for all assessed joints): MCPs I-V, wrist (radio-carpal, radio-ulnar and inter-carpal), elbow, knee, talo-crural and MTP I-V bilaterally. The decision of whether to adjust medication was based on change in and the level of the Disease Activity Score. If the patient did not respond, the treating physician immediately adjusted the therapy by proceeding to the next step in the treatment algorithm. If a patient responded or had reached the target, current medication was continued. In the ultrasound tight control group, the physician should overrule the decision based on the Disease Activity Score and proceed to the next step based on ultrasound findings. In both groups, clinically swollen joints were treated by intra-articular glucocorticoids when indicated. In the ultrasound tight control group an additional target was all joints with power Doppler signal, and all injections had to be ultrasound guided. For both groups, intra-articular injections of only tender joints were not allowed. The maximum dosage of triamcinolone hexacetonid per visit was 80 mg which could be distributed within joints as decided by the treating rheumatologist. . Duration 24 months. Concurrent medication/care: Participants in both groups were treated according to the

	same fixed treatment algorithm, adhering to a treat-to-target strategy with DMARD escalation therapy if target was not met. The treatment adjustments (including i.a. injections) that could be made were defined in a pre-specified dosing regimen. The initial treatment was methotrexate 15 mg/week increased to 20 mg/week by week five, in combination with seven weeks of prednisolone with tapering doses from 15 mg to zero. Further steps in the treatment algorithm included methotrexate 25 mg/week, triple synthetic disease modifying drug therapy (methotrexate, sulfasalazine, hydroxychloroquine) and biologic treatment according to guidelines. NSAIDs and coxibs were permitted. The choice and dosage of NSAIDs/coxibs was at the discretion of the treating rheumatologist. Analgesics up to the maximum recommended dose could be used for pain relief as required. Patients had to avoid analgesics within 24 hours prior to a visit if possible. All participants received vitamin D and calcium supplement during treatment with glucocorticoids ≥7.5mg, and postmenopausal women and older men (>70 years) was considered for a bisphosphonate according to general guidelines. IV or IM glucocorticoids were not allowed during the study. Oral glucocorticoids were allowed. Other DMARDs, besides methotrexate, sulfasalazine, hydroxychloroquine, biologics,and glucocorticoids, were not allowed. . Indirectness: No indirectness Comments: The decision of whether to adjust medication was based on change in and the level of the Disease Activity Score. US findings were used to decide on next steps in treatment and to guide intra-articular injections of glucocorticoids if required. (n=116) Intervention 2: Monitoring using clinical assessment alone (usual care) - Monitoring without ultrasound. Assessed at 0, 1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 20, and 24 months. Assessed by ultrasound yearly, but both the patient and the treating physician were blinded to the results, and the treating physicians did not have access privileges to ultrasound data in the electronic case report form. The treatment target in the conventional tight control strategy followed was clinical remission (Disease Activity Score <1.6 and no swollen joints). The decision of whether to adjust medication was based on change in and the level of the Disease Activity Score. If the patient did not respond, the therapy was adjusted by proceeding to the next step in the treatment algorithm. If a patient responded or had reached the target, current medication was continued. . Duration 24 months. Concurrent medication/care: Patients in both groups were treated according to the same fixed treatment algorithm, adhering to a treat-to-target strategy with DMARD escalation therapy if target was not met. The treatment adjustments (including i.a. injections) that could be made were defined in a pre-specified dosing regimen. The initial treatment was methotrexate 15 mg/week increased to 20 mg/week by week five, in combination with seven weeks of prednisolone with tapering doses from 15 mg to zero. Further steps in the treatment algorithm included methotrexate 25 mg/week, triple synthetic disease modifying drug therapy (methotrexate, sulfasalazine, hydroxychloroquine) and biologic treatment according

	to guidelines. NSAIDs and coxibs were permitted. The choice and dosage of NSAIDs/coxibs was at the discretion of the treating rheumatologist. Analgesics up to the maximum recommended dose could be used for pain relief as required. Patients had to avoid analgesics within 24 hours prior to a visit if possible. All patients received vitamin D and calcium supplement during treatment with glucocorticoids ≥ 7.5mg, and postmenopausal women and older men (>70 year) was considered for a bisphosphonate according to general guidelines. IV or IM glucocorticoids were not allowed during the study. Oral glucocorticoids were allowed. Other DMARDs, besides methotrexate, sulfasalazine, hydroxychloroquine, biologics, and glucocorticoids, were not allowed. In both groups, clinically swollen joints were treated by intra-articular glucocorticoids when indicated. For both groups, intra-articular injections of only tender joints were not allowed. The maximum dosage of triamcinolone hexacetonid per visit was 80 mg which could be distributed within joints as decided by the treating rheumatologist.. Indirectness: No indirectness
Funding	Academic or government funding (In addition grants were received from pharmaceutical industry.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ULTRASOUND MONITORING - JOINTS MIXED versus MONITORING WITHOUT ULTRASOUND

Protocol outcome 1: Disease activity score at 12 months
- Actual outcome: Change in DAS28 at 12 months; Group 1: mean -2.4 (SD 1.3); n=118, Group 2: mean -2.4 (SD 1.4); n=112; DAS28 2.0-10.0 Top=High is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR

Protocol outcome 2: Quality of life at 12 months
- Actual outcome: Change in EQ-5D at 12 months; Group 1: mean 0.28 (SD 0.28); n=118, Group 2: mean 0.25 (SD 0.29); n=112; EQ-5D -0.59-1 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR
- Actual outcome: Change in rheumatoid arthritis impact of disease score at 12 months; Group 1: mean -2.6 (SD 2); n=118, Group 2: mean -2.4 (SD 2.3);

n=112; Rheumatoid arthritis impact of disease score (RAID) 0-10 Top=High is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR

Protocol outcome 3: Remission at 12 months
- Actual outcome: DAS remission (DAS <1.6) at 12 months; Group 1: 76/118, Group 2: 81/112
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR

Protocol outcome 4: Pain at 12 months
- Actual outcome: Change in pain visual analogue scale at 12 months; Group 1: mean -32.5 (SD 24.8); n=118, Group 2: mean -29.2 (SD 28.1); n=112; Visual anlogue scale 0-100 Top=High is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR

Protocol outcome 5: Radiological progression at 12 months
- Actual outcome: Changes in modified Sharp score at 24 months; Median (IQR): US group: 1.0 (0-2.5) Usual care: 1.5 (0.5-3.0) modified Sharp score 0-448 Top=High is poor outcome;
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR

Protocol outcome 6: Withdrawal from trial / adherence to strategy at longest reported by study
- Actual outcome: Discontinuation due to no longer willing at 24 months; Group 1: 2/106, Group 2: 5/105
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention

group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR - Actual outcome: Discontinuation due to adverse event at 24 months; Group 1: 6/110, Group 2: 5/105 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Imputation was done for missing data. Logistic regression was used for continuous outcomes and in the case of binary outcomes the worst outcome was assumed and imputed. Missing data at 12 months not reported. Complete data available for 104 patients in intervention group, 100 in usual care group.; Indirectness of outcome: No indirectness ; Baseline details: More women were randomised to intervention group.; Group 1 Number missing: 18, Reason: NR; Group 2 Number missing: 16, Reason: NR	
Protocol outcomes not reported by the study	Function at 12 months; Low disease activity at 12 months; Fatigue at 12 months; Relapse at 12 months; Flare at 12 months; Change in planned management at over study duration

D.21 Prognostic studies

Reference	Boyesen 20115; Haavardsholm 2008[26]
Study type and analysis	Prospective cohort study Stepwise multiple logistic regression
Number of participants and characteristics	N = 84 (79 analysed) Country: Norway Prognostic factors (baseline) of 79 patients split by presence or absence of MRI erosive progression: Patients with MRI erosive progression (N=53): DAS28, median (IQR): 4.4 (3.0-5.4) USGS inflammation (<0.5 vs ≥0.5), median (IQR): 1.0 (1.0-2.0) Anti-CCP+, >25 U/ml, n (%): 31 (58.5) Total van der Heijde/Sharp score, median (IQR): 2.0 (0.0-6.0) Patients without MRI erosive progression (N=26): DAS28, median (IQR): 4.0 (3.3-4.9) USGS inflammation (<0.5 vs ≥0.5), median (IQR): 0.5 (0.0-1.0) Anti-CCP+, >25 U/ml, n (%): 13 (50.0) Total van der Heijde/Sharp score, median (IQR): 2.0 (0.0-4.0) Inclusion criteria: Patients with rheumatoid arthritis (according to ACR 1987) with a disease duration <1 year (date from diagnosis). Exclusion criteria: not mentioned Population characteristics (baseline) of 79 patients: Patients with MRI erosive progression (N=53): Female, n (%): 38 (71.7) Age, median (IQR): 57.5 (48.7-65.1) Disease duration, days, median (IQR): 112.0 (70.0-190.0) Immunoglobulin M (IgM) RF+, >25 U/ml, n (%): 26 (49.1) ESR, mm/h, median (IQR): 15.0 (7.5-25.0) CRP, mg/l, median (IQR): 5.2 (2.1-12.3)

Reference	Boyesen 20115; Haavardsholm 2008[26]
	Patients without MRI erosive progression (N=26): Female, n (%): 23 (88.5) Age, median (IQR): 58.7 (44.9-68.2) Disease duration, days, median (IQR): 91.5 (61.8-192.5) IgM RF+, >25 U/ml, n (%): 10 (38.5) ESR, mm/h, median (IQR): 12.0 (9.8-22.0) CRP, mg/l, median (IQR): 4.4 (1.9-7.6) Recruitment: 84 patients meeting the inclusion criteria were consecutively enrolled between February 2002 and June 2004. Seventy-nine patients completed the 1-year follow-up. Assessment: Visits were scheduled at baseline, 3, 6 and 12 months, and included 28 swollen and tender joint counts, patient reported health status by questionnaires, blood samples, laboratory analyses and imaging procedures (MRI, ultrasonography and conventional radiographs). The DAS28 was computed based on 28 joint counts and ESR. Ultrasound grey-scale (USGS) of 5 assessed locations in the dominant wrist was performed by a 'trained user' on a Diasus machine without power Doppler function. Only 70 (83%) patients had a baseline US examination due to logistical problems. DMARD use: Patients were treated according to clinical practice. DMARDs were used by 77.4% of the included patients (57.1% methotrexate monotherapy, 8.3% sulfasalazine monotherapy, 7.1% hydroxychloroquine monotherapy, 3.6% DMARD combination therapy). Anti-tumour necrosis factor α treatment was used by one patient (1.2%), and 60.7% of patients received glucocorticoids. At 1 year, follow-up DMARDs were used by 91.8% of patients, anti-tumour necrosis factor α drugs by two patients (2.6%) and glucocorticoids by 49.3%.
Prognostic variable(s)	USGS inflammation (<0.5 vs ≥0.5), DAS28, anti-CCP+, radiographic damage (van der Heijde/Sharp score)
Confounders	Univariate analysis considered above variable plus: age, sex, disease duration, ESR, CRP, IgM RF+, MRI synovitis, MRI bone marrow oedema (BME), MRI tenosynovitis, digital x-ray radiogrammetry bone mineral density (DXR BMD) g/m², DXR BMD (3 months change),
Outcomes and effect sizes	Outcome: MRI erosive progression (at 1 year) (53 (67%) patients experienced the outcome) Variables entered into model: age, sex, CRP, MRI synovitis, MRI BME, USGS inflammation, DXR BMD (3 months change). MRI erosive progression was defined as a one or more unit increase in 1-year rheumatology magnetic resonance imaging scoring system (RAMRIS) erosive change. Possible associations between imaging modalities and the outcome were explored by logistic

Reference	Boyesen 20115; Haavardsholm 2008[26]
	regression analyses. Age, sex and independent variables with univariate associations of p≤0.25 were included in the multivariate analyses. The final multivariate model with independent predictors of 1-year change in MRI erosions was obtained by stepwise exclusion of the least significant variable until only significant variables were left. The USGS inflammation score was calculated in the following way: the scanned tendon areas were assessed for USGS tenosynovitis defined as hypoechoic, poorly compressible, thickened tendon. All findings were graded as 0=none, 1=mild, 2=moderate or 3=marked. The 5 assessed locations were dichotomised into present versus not present and summed into a USGS inflammation score ranging from 0 to 5. Cut-off point for presence or absence of USGS inflammation was set at 0.5. Results: Final model included MRI BME, sex and age. USGS inflammation: OR 2.01 (95% CI 1.14-3.53) The other prognostic factors (DAS28, anti-CCP+, van der Heijde/Sharp score) were not independently associated with the outcome of MRI erosive progression.
Comments	Low risk of bias

1

2

Reference	Geng 2016[23]
Study type and analysis	Prospective observational study Multivariable logistic regression
Number of participants and characteristics	N = 126 Country: China Prognostic factors (baseline) of 126 patients: DAS28-ESR, mean (SD):1.89 (0.50) DAS28-CRP, mean (SD) 1.41 (0.38) PD > 0, n (%): 61 (48%) SH > 0, n (%): 73 (58%) PD total score, mean (SD): 2.0 (3.5) SH total score, mean (SD): 2.6 (3.6) Ultrasonographic remission (PD = 0 and SH = 0), n (%): 53 (42%)

Reference	Geng 2016[23]
	Tenosynovitis, n (%): 25 (20%) Bone erosion, n (%): 55 (44%). Inclusion criteria: Patients in clinical remission (DAS28-ESR ≤ 2.6 at two consecutive visits 3 months apart), fulfilling the 2010 ACR/EULAR classification criteria for diagnosis of RA, RA treatment stable for at least 3 months and no clinical indication for a change in treatment. Exclusion criteria: None reported. Population characteristics (baseline) of 126 patients: Female, n (%): 92 (73%) Age, mean (SD): 48.4 (15.3) Disease duration, years, mean (SD): 5.2 (6.1) RF+, n (%): 93 (74%) Anti-CCP+, n (%): 98 (78%) Recruitment: Patients attending the rheumatology clinic of Peking University First Hospital between March 2012 and June 2014 were enrolled. Assessment: Clinical and laboratory examinations, disease activity assessments and ultrasonography were performed every 3 months for each patient. Ultrasonography was performed by 2 well-experienced rheumatologists who were blinded to all clinical findings. 22 joints (bilateral wrists, MCP1-5) and PIP-(proximal interphalangeal joints 1-5) were scanned from dorsal aspect on transverse and longitudinal planes. MCP2 and MCP5 joints were additionally assessed from the lateral aspect. Each scan took at least 15 minutes. The Esoate Mylab 90 machine with a 6-18 Mhz transducer was used. Power Doppler (PD) subclinical synovitis and grey-scale synovial hypertrophy (SH) were measured and graded using the 2001 Sukudlarek semi-quantitative method on a scale of 0-3. PD and SH total scores were defined as the sum of the respective scores at each joint (0-66). Tenosynovitis and bone erosion were defined according to pathological changes in articular inflammatory diseases in OMERACT. DMARD use at baseline: Not reported (RA treatment stable for at least 3 months and no clinical indication for a change in treatment.)
Prognostic variable(s)	DAS28-ESR, PD>0, PD total score and SH total score

Reference	Geng 2016[23]
Confounders	Additional variables included in univariate analysis not known (assume all baseline variables measured were considered in the univariate analysis but data not shown).
Outcomes and effect sizes	Outcome: Relapse (DAS28-ESR > 2.6 following a period of clinical remission) [at 12 months]. 54 patients (43%) experienced the outcome. Variables entered into model: At least the above prognostic variables, whether any additional variables were entered into the multivariate model is unclear. Results: PD > 0, n (%): OR 8.8 (95% CI 2.7 – 28.4) PD total score: OR 1.4 (95% CI 0.9 – 2.0) SH total score: OR 0.7 (95% CI 0.5 – 1.0) DAS28-ESR was also independently associated with the outcome but the data has not been presented here as it was incorrectly reported by the authors.
Comments	Very high risk of bias (outcome measurement – blinding of outcome assessors and inter-rater measurement not reported; statistical analysis and reporting – UVA analysis not reported and MVA model unclear).

1

2

Reference	Horton 2016[32]
Study type and analysis	Prospective observational study Multivariable logistic regression
Number of participants and characteristics	N = 217 (105 analysed) Country: UK Prognostic factors (baseline) of 105 patients: DAS28-CRP3v, median (IQR): 4.5 (3.8-5.2) DAS28-CRP4v, median (IQR): 4.9 (4.0-5.5) DAS44-CRP4v, median (IQR): 3.1 (2.6-3.6) US of 26 joints Total greyscale (GS) score, median (IQR): 17 (10-25)

Reference	Horton 2016[32]
	Total power Doppler activity (PDA) score, median (IQR): 3 (0-8) Absence of PDA (total PDA=0), n (%): 23 (27) Absent/minimal PDA (total PDA≤1), n (%): 34 (40) Inclusion criteria: Patients enrolled between June 2010 and September 2012, fulfilling the ACR 1987 and/or 2010 ACR/EULAR criteria for the classification of RA, and DAS28-CRP4v ≥2.6 at baseline. Inclusion of patients with coexisting osteoarthritis (OA) was permitted due to the prevalence of these findings. Exclusion criteria: Patients not receiving DMARDs within 3 months of baseline (e.g. due to contraindications), receiving alternative non-RA diagnosis within the following 12 months or those with missing 12 month outcome data. Population characteristics (baseline) of 105 patients: Female, n (%): 79 (75) Age, mean (SD): 59 (13) Symptom duration, months, median (IQR): 6.0 (4.0-13.0) RF+, n (%): 78 (74) ACPA+, n (%): 81 (77) CRP, mg/l, median (IQR): 21 (7-45) HAQ, median (IQR): 1.3 (0.8-1.9) Radiographic erosion in the hands and feet, n (%): Any: 18 (17) 1987 ACR definition: 11 (10) 2010 ACR/EULAR definition: 9 (9) Recruitment: Patients with new-onset inflammatory arthritis attending the Leeds Early Arthritis Clinic. Assessment: Clinical data was collected every 3 months, or as clinically indicated, in accordance with EULAR guidelines. Assessments included examination of 44 joints for swelling and 53 joints for tenderness (including RAI) by rheumatologists and rheumatology nurse-specialists. All patients were managed according to the EULAR treat-to-target recommendations when clinically appropriate. The target was remission defined by DAS28 using 4 variables (DAS28-CRP4v<2.6): SJC28, TJC28, CRP and patient visual analogue scale disease assessment (VASDA). As this definition of remission may allow persistence of swollen joints, consultant impression of disease remission also factored in treatment decisions, in accordance with guidelines. Treatment escalation to biologic therapy was as recommended by NICE, that is at least high disease activity (DAS28>5.1) after failure of at least two

Reference	Horton 2016[32]
	synthetic DMARDs including methotrexate. Ultrasound examination of 26 joints was performed at baseline and at 12 months. A lower number of joints were assessed by US than clinically in order to 'optimise feasibility'. US was carried out in a routine outpatient setting using a GE E9 machine by a validated sonographer who had been trained by an experienced EULAR teacher. DMARD use at baseline: Most patients commenced DMARDs at baseline (72%) or within 4 weeks (15%). The first DMARD chosen was MTX in 86%, HCQ in 10% and SSZ in 4%. MTX in combination was commenced in the remaining 1%. Over 12 months 9% of patients failed MTX and 19% failed another DMARD. By 12 months, 24% of patients were on MTX plus another DMARD. 92% of patients were assessed every 3 months until the target for treatment (DAS28 < 2.6) was met.
Prognostic variable(s)	TJC28, SJC28, CRP mg/L, patient VAS global disease assessment (components of DAS28-CRP), total GS score on US, total PDA score on US
Confounders	Univariate analysis considered above variable plus: age, sex, BMP, symptom duration, smoking status, number of comorbidities, concurrent OA, RF+, ACPA+, HAQ, radiographic erosions.
Outcomes and effect sizes	Outcome: Remission: both DAS28-CRP4v < 2.6 and DAS44-CRP4v < 1.6 [at 12 months]. 45 patients (43%) and 41 patients (39%) experienced the two outcomes respectively. Variables entered into model: Variables demonstrating statistical significance ($p < 0.05$) on univariate analysis (symptom duration, TJC28, patient global disease assessment (DAS28 remission outcome only), HAQ, female sex (DS44 remission outcome only)). DAS-28CRP3v was also significant on univariate analysis but was not entered in to the multivariable analysis because of overlap with its component variables. Results: DAS28-CRP4v < 2.6 TJC28 OR 0.93 (95% CI 0.85 – 1.02) Patient VAS global disease assessment OR 0.98 (95% CI 0.95 – 1.00) Final model also included symptom duration ($p = 0.04$) and HAQ (NS). Results: DAS44-CRP4v < 1.6 TJC28 OR 0.88 (95% CI 0.79 – 0.98) Final model also included female sex ($p = 0.02$), symptom duration (NS) and HAQ (NS).
Comments	Very high risk of bias (study attrition – 52% patients lost to follow up or excluded for incomplete outcome data; outcome measurement – blinding of outcome assessors and inter-rater measurement not reported; statistical analysis and reporting – only included variables in MVA if $p < 0.05$ in UVA, may have missed important variables).

1

2

Reference	Saleem 2012[60]
Study type and analysis	Cohort study Multiple binary logistic regression
Number of participants and characteristics	N = 93 (83 analysed) Country: UK Prognostic factors (baseline) of 93 patients: DAS28, median (IQR): 2.29 (1.79-3.19) US GS synovial hypertrophy, n (%): 83 (89.2) US PD activity, n (%): 58 (62.4) US erosions, n (%): 65 (69.9) Inclusion criteria: Patients fulfilling the ACR 1987 criteria for the classification of RA, aged >18 years, without flares of disease in the last 6 months, stable treatment for 6 months and no indication for a change in treatment. Exclusion criteria: not mentioned Population characteristics (baseline) of 93 patients: Female, n (%): 63 (67.7) Age, mean (95% CI): 56.6 (53.9-59.4) Disease duration, years, median (IQR): 7.0 (4.5-9.5) Remission duration, months, median (IQR): 22 (12-34) RF+, n (%): 39 (41.9) Anti-CCP+, n (%): 38 (59.4) CRP>0, n (%): 51 (54.8) CRP, median (IQR): 5 (0-8) HAQ-DI score, mean (95% CI): 0.337 (0.277-0.397) RAQoL, mean (95% CI): 6.4 (5.3-7.6) Simplified Disease Activity Index (SDAI) remission, n (%): 31 (33.3) DAS28 remission, n (%): 52 (55.9)

Reference	Saleem 2012[60]
	1981 ACR remission, n (%): 50 (53.8) 2011 ACR/EULAR remission, n (%): 13 (14.0) Recruitment: Consecutive RA patients deemed to be in clinical remission by their assessing consultant rheumatologist were recruited. All patients were taking conventional DMARDs. Leeds Teaching Hospitals NHS trust provided ethics approval. Assessment: Flare was defined as any increase in disease activity that required an initiation, change or increase in therapy. This was assessed clinically and biochemically. Clinical assessments were performed at baseline, every 3 months and at the time of flare over a period of 1 year. The assessments included duration of morning stiffness, global assessment of health and disease activity, quality of life, number of tender and swollen joints. Laboratory tests for CRP, anti-CCP and RF were performed. Ultrasound was performed by a single experienced ultrasonographer using a Phillips ATL HDI 3000 machine. DMARD use at baseline: methotrexate: 40%, sulphasalazine: 24%, leflunomide: 4%, hydroxychloroquine: 2%, gold salts: 2%, combination dual therapy: 13%, combination triple therapy: 5%, no DMARDs: 10%
Prognostic variable(s)	Power Doppler (PD) present, remission (DAS28<2.6)
Confounders	HAQ-DI
Outcomes and effect sizes	Outcome: disease flare (defined as any increase in disease activity, using DAS28 criteria that required an initiation, change or increase in therapy. This was assessed clinically and biochemically.) At 1 year, 24 patients (26%) experienced the outcome. Variables entered into model: the prognostic and confounder variables Binary logistic regression was used to identify associations between the various clinical and imaging assessments at baseline and the odds of subsequent flare. 'Given the number of patients who flared was relatively small (26%) it was not possible to construct a fully comprehensible multiple logistic regression model to identify factors that were independently associated with the odds of flare. Only PD absent/present and HAQ-DI score in combination with the various remission criteria were included. RAQoL was highly correlated with HAQ-DI (Pearson's r=0.81) therefore it was not included in the multivariable model.' Different MVA models were run for the different remission criteria. Only the results of the DAS28 model are reported below. Results: PD present: OR 7.57 (95% CI 1.75-32.76) Remission (DAS28<2.6): OR 2.71 (95% CI 0.73-10.14) Final model also included HAQ-DI.

Reference	Saleem 2012[60]
Comments	Very high risk of bias (study attrition- lacking information on number of drop-outs and how it was accounted for in analysis; prognostic factor measurement- authors chose crude measurement of DAS28 remission (dichotomous) over continuous DAS; study confounding- few possible confounders accounted for in MVA model due to small event numbers)

1

2

Reference	Zavada 2017[76]
Study type and analysis	Prospective cohort study Multiple logistic regression
Number of participants and characteristics	N = 185 (185 analysed) Country: Czech Republic Prognostic factors (baseline) of 185 patients: DAS28-CRP, mean (SD): 3.7 (1.5) Greyscale synovitis sum score (GSsynSS), mean (SD): 6.91 (6.39) Power Doppler synovitis sum score (PDsynSS), mean (SD): 4.02 (5.20) Greyscale tenosynovitis sum score (GStenSS), mean (SD): 0.66 (1.15) Power Doppler tenosynovitis sum score (PDtenSS), mean (SD): 0.68 (1.63) US erosion score (ES), mean (SD): 1.18 (2.10) Inclusion criteria: Patients (N=46; incident) fulfilling the ACR/EULAR 2010 criteria for the classification of RA with either early RA newly started on therapy with conventional synthetic disease-modifying drugs (csDMARDs) or glucocorticoids, and patients (N=139; prevalent) with established RA. Exclusion criteria: Patients on biological DMARDs as they are followed separately in the Czech biologics ATTRA registry. Population characteristics (baseline) of 185 patients: Female, n (%): 142 (76.8) Age, mean (SD): 55.2 (14.0) Disease duration, years, mean (SD): 6.3 (7.7)

Reference	Zavada 2017[76]
	HAQ, mean (SD): 0.77 (0.73) RF+, n (%): 87 (47.2) ACPA+, n (%): 116 (62.9) CRP, mg/l, mean (SD): 7.6 (9.0) Recruitment: All patients were recruited from the outpatient rheumatology clinic at the Institute of Rheumatology in Prague and were followed longitudinally according to a predefined protocol. During the observation period, patients were routinely treated by their rheumatologist. Assessment: The same measures of clinical assessment and US imaging were applied at baseline and every 6 months during the follow-up period. Physical function was assessed by the HAQ at baseline and then annually over a three year period. For the purpose of this analysis, only clinical and US data collected concurrently with the HAQ every 12 months were used. Tender and swollen joints counts were carried out on 28 joints in accordance with EULAR recommendations. At baseline, demographic and anthropometric data was collected and RF and ACPA testing performed. The US examinations were performed using Esaote Mylab 60 by 8 clinicians who had undergone intermediate to advanced US training. DMARD use at baseline: oral glucocorticoids: 43.8%, csDMARDs: 87.0%
Prognostic variable(s)	Previous (12 months earlier) DAS28-CRP and previous US7 assessment (GSsynSS, PDsynSS, GStenSS, PDtenSS, ES)
Confounders	Univariate analysis considered above variables plus: previous HAQ
Outcomes and effect sizes	Outcome: HAQ score [12 months after prognostic variables measured] Variables entered into model: all of the prognostic variables and the confounder variable mentioned above plus sex, age, body mass index (BMI), RF+ or ACPA+, prevalent vs. incident RA which had not been tested for associations in univariate analyses A linear mixed effects model was used to study the longitudinal relationship between the clinical measures of disease activity (DAS28-CRP), US inflammatory score (German US7 scores sum-scores for GS and PD synovitis, GS and PD tenosynovitis, and erosions) as explanatory variables, and physical function (HAQ score) as a dependent variable. To explore the impact of previous disease activity (assessed either clinically or by US) On the current HAQ score, the authors applied a 'time-lag model, which related the covariates measured at the previous visit (12 months before) to the outcome variable assessed at the current visit'. Also, the previous HAQ was added to the model (i.e. first-order autoregression) to model change in HAQ score rather than absolute HAQ scores. No step-wise elimination was performed. Statistical significance was set at p≤0.05.

Reference	Zavada 2017[76]
	Results: Results were reported for all variables entered into the multivariate model. Previous DAS28-CRP: β coefficient 0.161 (95% CI 0.113 to 0.208) Previous GSsynSS: β coefficient -0.004 (95% CI -0.019 to 0.010) Previous PDsynSS: β coefficient -0.021 (95% CI -0.040 to -0.002) Previous GStenSS: β coefficient 0.000 (95% CI -0.085 to 0.085) Previous PDtenSS: β coefficient -0.015 (95% CI -0.078 to 0.048) Previous ES: β coefficient 0.012 (95% CI -0.022 to 0.046)
Comments	High risk of bias (study attrition- missing data unclear) Serious indirectness (MVA model is not limited to impact of baseline variables on outcome at follow up, but rather looks at the association between prognostic variables at multiple time points to the outcome measured 12 months later (i.e. association between month 0 variables and month 12 outcome, month 12 variables and month 24 outcome, month 24 variables and month 36 outcome, all in one analysis).

Appendix E: Forest plots

E.1 Randomised controlled trial: Clinical assessment and ultrasound versus clinical assessment alone

Figure 2: Change in Disease activity score (DAS28)

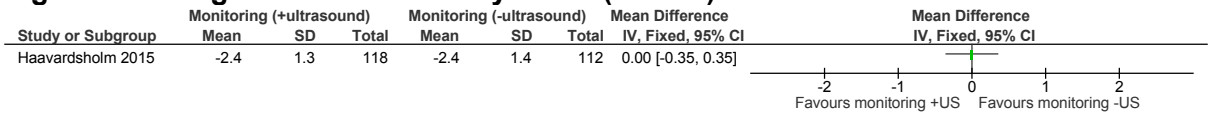


Figure 3: Change in function (RAID score)

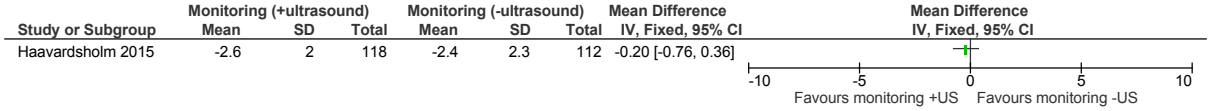


Figure 4: Change in Quality of life (EQ-5D)

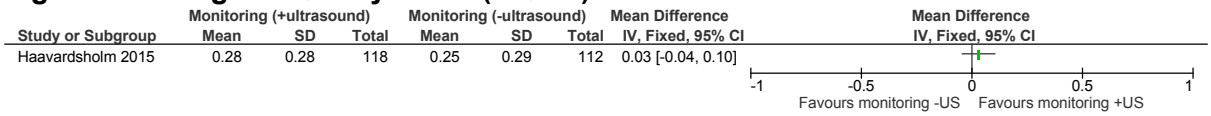


Figure 5: Remission (DAS<1.6)

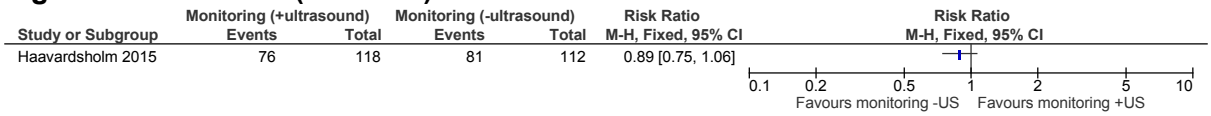


Figure 6: Change in pain (VAS)

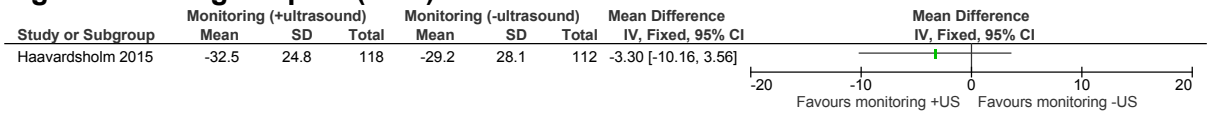


Figure 7: Withdrawal from trial due to "no longer willing"

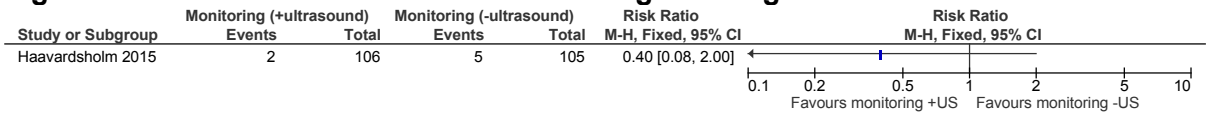
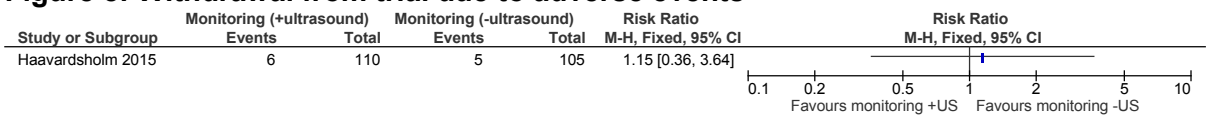


Figure 8: Withdrawal from trial due to adverse events



E.2 Prognostic studies: Effect of clinical and US prognostic factors on various outcomes

Note: All prognostic factors are displayed on the forest plots even where odds ratios were not reported, as these factors were considered by the studies. Where a study has its results listed as 'Not estimable' for a specific factor, that factor was not independently associated with the outcome following multivariable analysis.

Figure 9: Clinical and US prognostic factors for MRI erosive progression (dichotomous; at 1 year)

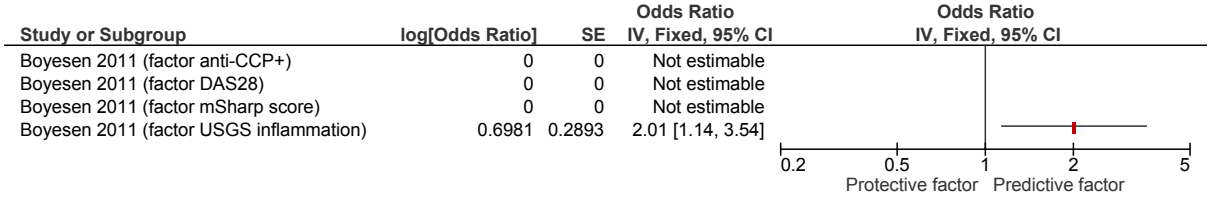


Figure 10: Clinical and US prognostic factors for disease flare (increase in disease activity requiring an initiation, change or increase in therapy based on DAS28; dichotomous – at 1 year)

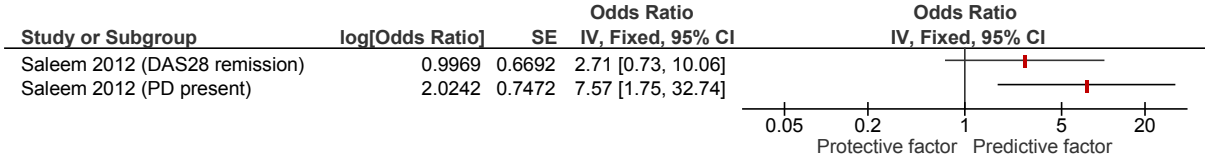
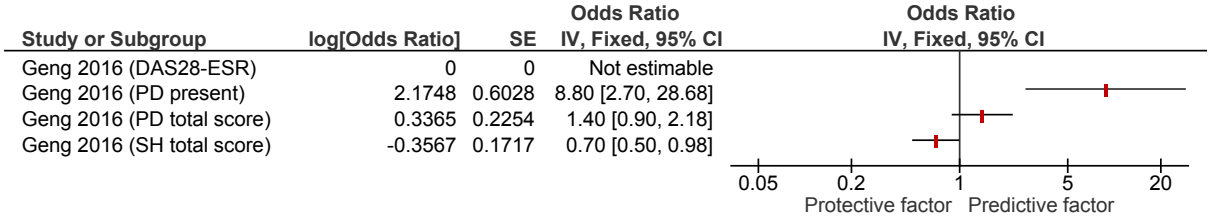


Figure 11: Clinical and US prognostic factors for relapse (DAS28-ESR>2.6 following a period of clinical remission; dichotomous – at 1 year)



Note: DAS28-ESR was independently associated with the outcome but data is not presented here as it was incorrectly reported by the authors.

Figure 12: Clinical and US prognostic factors for remission (DAS28-CRP4v <2.6; dichotomous – at1 year)

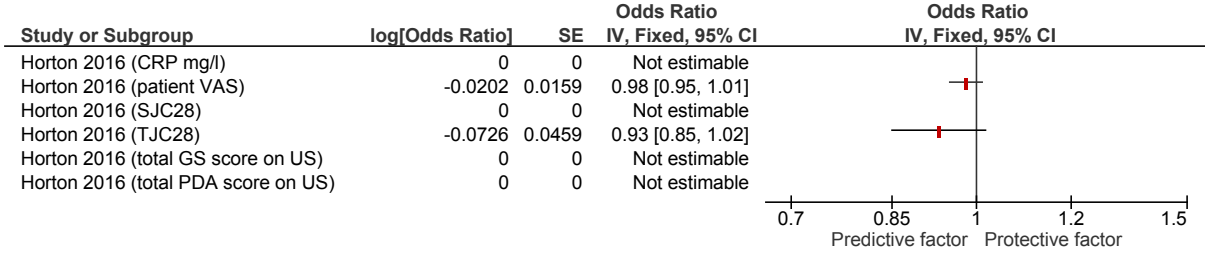
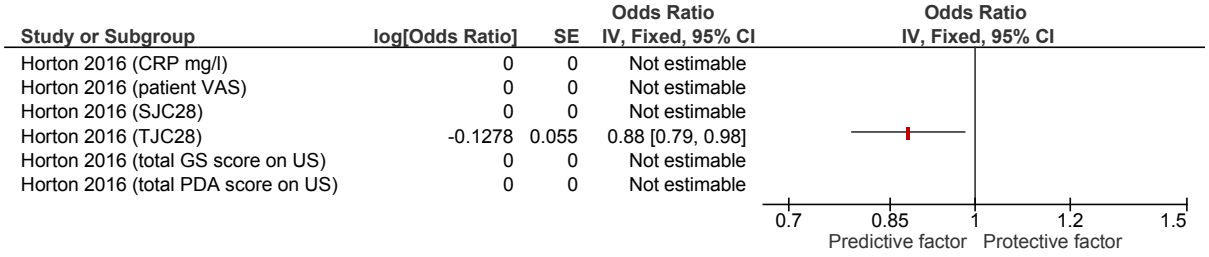


Figure 13: Clinical and US prognostic factors for remission (DAS44-CRP4v <1.6; dichotomous – at 1 year)



1

2

3

Appendix F: GRADE tables

F.1 Randomised controlled trials

Table 18: Clinical evidence profile: Clinical assessment and ultrasound versus clinical assessment alone

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Monitoring including ultrasound	Control	Relative (95% CI)	Absolute		
Change in disease activity score (follow-up 12 months; measured with: DAS28; range of scores: 2-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	118	112	-	MD 0 higher (0.35 lower to 0.35 higher)	⊕⊕⊕O MODERATE	CRITICAL
Change in rheumatoid arthritis impact of disease score (follow-up 12 months; measured with: RAID score; range of scores: 0-10; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	118	112	-	MD 0.2 lower (0.76 lower to 0.36 higher)	⊕⊕⊕O MODERATE	CRITICAL
Change in quality of life (follow-up 12 months; measured with: EQ-5D; range of scores: -0.59-1; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	118	112	-	MD 0.03 higher (0.04 lower to 0.1 higher)	⊕⊕⊕O MODERATE	CRITICAL
Remission (follow-up 12 months; assessed with: DAS <1.6)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	76/118 (64.4%)	81/112 (72.3%)	RR 0.89 (0.75 to 1.06)	80 fewer per 1000 (from 181 fewer to 43 more)	⊕⊕OO LOW	IMPORTANT
Change in pain (follow-up 12 months; measured with: visual analogue scale; range of scores: 0-100; Better indicated by lower values)												
1	randomised	serious[1]	no serious	no serious	no serious	none	118	112	-	MD 3.3 lower (10.16 lower to	⊕⊕⊕O	IMPORTANT

	trials		inconsistency	indirectness	imprecision					3.56 higher)	MODERATE	
Change in radiological progression (follow-up 24 months; measured with: Sharp score; range of scores: 0-448; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	not assessed[3]	none	118	112	-	The median (IQR) change in: Control group: 1.5 (0.5-3.0), Intervention group: 1.0 (0-2.5)	⊕⊕⊕⊕ MODERATE	IMPORTANT
Withdrawal from trial due to "no longer willing" (follow-up 24 months)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	2/106 (1.9%)	5/105 (4.8%)	RR 0.4 (0.08 to 2)	29 fewer per 1000 (from 44 fewer to 48 more)	⊕OOO VERY LOW	IMPORTANT
Withdrawal from trial due to adverse events (follow-up 24 months)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	6/110 (5.5%)	5/105 (4.8%)	RR 1.15 (0.36 to 3.64)	7 more per 1000 (from 30 fewer to 126 more)	⊕OOO VERY LOW	IMPORTANT

1 [1] Downgraded by 1 increment because the majority of the evidence was at high risk of bias
2 [2] Downgraded by 1 increment because the confidence interval crossed one MID and downgraded by 2 increments if the confidence interval crossed both MIDs.
3 [3] Imprecision could not be assessed because non-parametric statistics were used. The confidence interval is relatively wide.

F.24 Prognostic studies

5 **Table 19: Clinical evidence profile: MRI erosive progression (dichotomous - at 1 year)**

Quality assessment							Effect	Quality
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations (including publication bias where possible)**	**Pooled effect (95% CI)**	
USGS inflammation								
1	Cohort studies	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	Adjusted OR: 2.01 (1.14-3.53)	HIGH
DAS28								

1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
Anti-CCP+								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
van der Heijde/Sharp score								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a

1 n/a: unable to assess as data not reported (factor not independently associated with the outcome following multivariable analysis)

2 **Table 20: Clinical evidence profile: Disease flare (dichotomous – at 1 year; increase in disease activity requiring an initiation, change**
3 **or increase in therapy based on DAS28)**

Quality assessment							Effect	Quality
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations (including publication bias where possible)**	**Pooled effect (95% CI)**	
Remission (DAS28<2.6)								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	Adjusted OR: 2.71 (0.73-10.14)	VERY LOW
PD present								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	Adjusted OR: 7.57 (1.75-32.76)	LOW

4 [1] Downgraded by 2 increments because the majority of the evidence was at very high risk of bias
5 [2] Downgraded by 1 increment because the confidence interval crosses the line of no effect

6

Table 21: Clinical evidence profile: Relapse (dichotomous – at 1 year; DAS28-ESR>2.6 following a period of clinical remission)

Quality assessment							Effect	Quality
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (including publication bias where possible)	Pooled effect (95% CI)	
PD>0								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	Adjusted OR: 8.8 (2.7 – 28.4)	LOW
PD total score								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	Adjusted OR: 1.4 (0.9 – 2.0)	VERY LOW
SH total score								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	Adjusted OR: 0.7 (0.5 – 1.0)	VERY LOW
DAS28-ESR								
1	Cohort studies	not assessed	not assessed	not assessed	not assessed	not assessed	data unavailable[3]	not assessed

[1] Downgraded by 2 increments because the majority of the evidence was very high risk of bias
[2] Downgraded by 1 increment because the confidence interval crosses the line of no effect
[3] Variable was independently associated with the outcome but data is not presented here as it was incorrectly reported by the authors.

Table 22: Clinical evidence profile: Remission (dichotomous – at 12 months; DAS28-CRP4v <2.6)

Quality assessment							Effect	Quality
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (including publication bias where possible)	Pooled effect (95% CI)	

CRP mg/l								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
Patient VAS global disease assessment								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	Adjusted OR: 0.98 (0.95 – 1.00)	VERY LOW
SJC28								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
TJC28								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	Adjusted OR: 0.93 (0.85 – 1.02)	VERY LOW
Total GS score on US								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
Total PDA score on US								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a

[1] Downgraded by 2 increments because the majority of the evidence was very high risk of bias
[2] Downgraded by 1 increment because the confidence interval crosses the line of no effect
n/a: unable to assess as data not reported (factor not independently associated with the outcome following multivariable analysis)

Table 23: Clinical evidence profile: Remission (dichotomous – at 12 months; DAS44-CRP4v <1.6)

Quality assessment							Effect	Quality
Number of	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Pooled effect**	

studies						(including publication bias where possible)	(95% CI)	
CRP mg/l								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
Patient VAS global disease assessment								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
SJC28								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
TJC28								
1	Cohort studies	very serious risk of bias[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	Adjusted OR: 0.88 (0.79 – 0.98)	LOW
Total GS score on US								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a
Total PDA score on US								
1	Cohort studies	n/a	n/a	n/a	n/a	n/a	Not independently associated with the outcome following multivariable analysis.	n/a

[1] Downgraded by 2 increments because the majority of the evidence was very high risk of bias
n/a: unable to assess as data not reported (factor not independently associated with the outcome following multivariable analysis)

Table 24: Clinical evidence profile: Function (continuous – HAQ score 12 months after prognostic variables were measured)

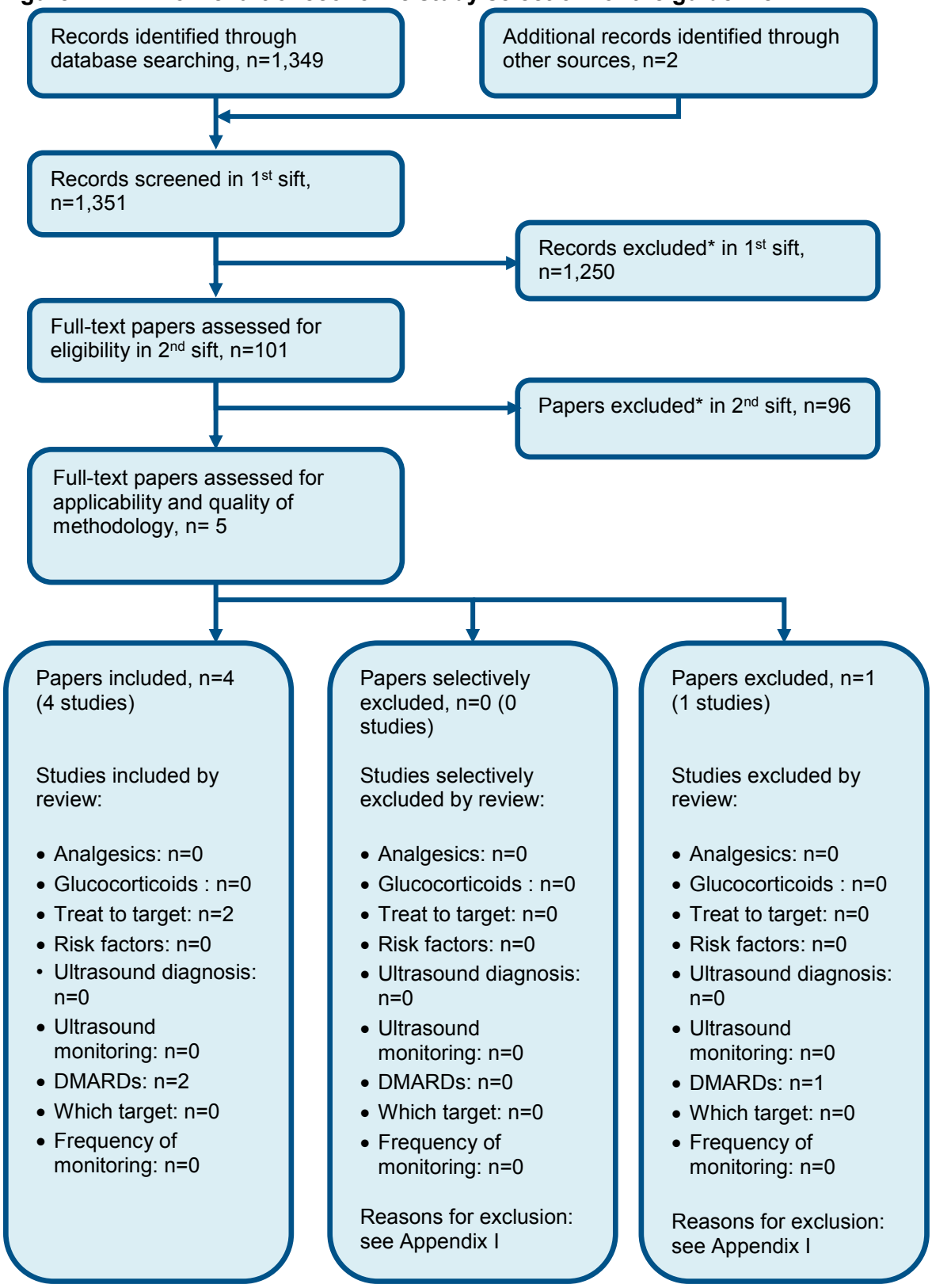
Quality assessment	Effect	Quality

Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations (including publication bias where possible)	Pooled effect (95% CI)	
Previous DAS28-CRP								
1	Cohort studies	serious risk of bias[1]	no serious inconsistency	serious indirectness[2]	no serious imprecision	none	Coefficient: 0.161 (0.113 to 0.208)	LOW
Previous GSsynSS								
1	Cohort studies	serious risk of bias[1]	no serious inconsistency	serious indirectness[2]	serious imprecision[3]	none	Coefficient: -0.004 (-0.019 to 0.010)	VERY LOW
Previous PDsynSS								
1	Cohort studies	serious risk of bias[1]	no serious inconsistency	serious indirectness[2]	no serious imprecision	none	Coefficient: -0.021 (-0.040 to -0.002)	LOW
Previous GStenSS								
1	Cohort studies	serious risk of bias[1]	no serious inconsistency	serious indirectness[2]	serious imprecision[3]	none	Coefficient: 0.000 (-0.085 to 0.085)	VERY LOW
Previous PDtenSS								
1	Cohort studies	serious risk of bias[1]	no serious inconsistency	serious indirectness[2]	serious imprecision[3]	none	Coefficient: -0.015 (-0.078 to 0.048)	VERY LOW
Previous erosion score								
1	Cohort studies	serious risk of bias[1]	no serious inconsistency	serious indirectness[2]	serious imprecision[3]	none	Coefficient: 0.012 (-0.022 to 0.046)	VERY LOW

[1] Downgraded by 1 increment because the majority of the evidence was high risk of bias
[2] Downgraded by 1 increment because the MVA model was not limited to impact of baseline variables on outcome at follow up, but rather looks at the association between prognostic variables at multiple time points to the outcome measured 12 months later
[3] Downgraded by 1 increment because the confidence interval crosses the line of no effect

Appendix G: Health economic evidence selection

Figure 14: Flow chart of economic study selection for the guideline



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

1

Appendix H: Health economic evidence tables

None.

1 Appendix I: Excluded studies

I.1 2 Excluded clinical studies

3 **Table 25: Studies excluded from the clinical review**

Reference	Reason for exclusion
Aydin 2017[1]	Unobtainable
Backhaus 2013[2]	Analyses not adjusted for confounders
Bellis 2016[4]	cross-sectional study, no follow-up
Brown 2008[6]	Analyses not adjusted for confounders
Bugatti 2016[7]	no relevant outcomes
Bugatti 2012[8]	Analyses not adjusted for all key confounders; outcome is indirect
Bugatti 2012[50]	see above
Cavet 2009[9]	Unclear if analyses are adjusted for confounders; incomplete reporting of results
Chen 2017[10]	Analyses not adjusted for key confounders
Cheung 2016[12]	Key confounders for the outcome (radiographic progression) not considered in analysis
Cheung 2014[11]	no regression analysis; no anti-CCP measured
Christensen 2016[13]	Study duration only 4 months
D'Agostino 2016[14]	Analyses not adjusted for confounders
D'Agostino 2017[15]	narrative review
Dale 2016[17]	mixed study population of RA and UA (breakdown unknown)
Dale 2014[16]	mixed study population of RA and UA (breakdown unknown)
Dougados 2012[18]	no regression analysis; no anti-CCP measured
Dougados 2013[19]	no regression analysis; no anti-CCP measured
El Miedany 2016[20]	results incompletely reported
Fukae 2017[21]	Analyses not adjusted for key confounders
Gartner 2015[22]	analyses not adjusted for confounders
Hammer 2017[27]	no regression analysis
Han 2016[28]	systematic review
Harman 2015[30]	Analyses not adjusted for confounders
Harman 2015[29]	multivariate analysis model does not include factors of interest
Hirata 2017[31]	Incorrect study design
Horton 2017[33]	No relevant outcomes
Hurnakova 2015[35]	cross-sectional study
Hurnakova 2016[34]	Incorrect study design
Iagnocco 2015[36]	Analyses not adjusted for key confounders; outcomes incompletely reported
Ikeda 2007[37]	Analyses not adjusted for confounders
Ikeda 2013[39]	Study looked at ultrasound alone not in combination with clinical assessment
Ikeda 2012[38]	see above
Ivanac 2015[40]	Study looked at ultrasound alone not in combination with clinical assessment
Janta 2016[41]	Analyses not adjusted for key confounders; results incomplete

Reference	Reason for exclusion
Jeka 2017[42]	no regression analysis
Jindal 2017[43]	Unobtainable
Kakavouli 2015[44]	single case study
Kamel 2017[45]	No relevant outcomes
Kawashiri 2017[46]	retrospective study
Kirino 2015[47]	retrospective study
Komarova 2015[48]	Incorrect study design
Luz 2016[49]	Analyses not adjusted for key confounders
Manzo 2012[50]	shouldn't have been ordered - exclude from EXCLUSION TABLE
Naredo 2007[51]	Analyses not adjusted for confounders
Naredo 2008[52]	no anti-CCP and no baseline erosions measured
Osipyants 2013[54]	not enough information on analyses (conference abstract)
Osipyants 2013[55]	not enough information on analyses (conference abstract)
Ramírez García 2014[56]	Adjustment for confounders unclear; incomplete reporting of results
Rees 2007[57]	no regression analysis
Reynolds 2009[58]	no regression analysis
Rosa 2016[59]	cross-sectional study; analyses not adjusted for key confounders
Scire 2009[61]	mixed patient population (either RA or undifferentiated polyarthritis)
Sreerangaiah 2016[62]	results incompletely presented; unclear if analyses was adjusted for key confounders
Takase-Minegishi 2017[63]	Systematic review: references checked
Tan 2016[64]	no regression analysis
Taylor 2004[65]	Unclear if analyses are adjusted for confounders; incomplete reporting of results
Tokai 2015[66]	Analyses not adjusted for key confounders
Toyota 2016[67]	Unobtainable
Valor 2016[68]	Incorrect study design
Van der ven 2017[69]	No relevant outcomes
Vlad 2015[70]	Analyses not adjusted for key confounders
Vreju 2016[71]	Key confounders for the outcome not considered in analysis
Wakefield 2007[72]	Analyses not adjusted for confounders
Yamada 2016[73]	Unobtainable
Yoshimi 2013[75]	no regression analysis; no anti-CCP measured
Yoshimi 2014[74]	no regression analysis; no anti-CCP measured
Zhao 2017[77]	Literature review

I.2 Excluded health economic studies

Table 26: Studies excluded from the health economic review

Reference	Reason for exclusion
None	

Appendix J: Research recommendations

J.1 Ultrasound to assess disease activity (monitoring) where clinical examination is inconsistent or inconclusive

Research question: What is the clinical and cost effectiveness of using ultrasound to monitor disease in adults with RA when/where clinical examination is inconclusive or inconsistent with other signs of disease activity?

Why this is important:

Rheumatoid arthritis is a chronic inflammatory condition which requires regular review of disease activity to enable relevant adjustments in management accordingly to achieve a target of remission or low disease activity.

While some people in clinical remission have been found to have subclinical inflammation or erosions on ultrasound examination, randomised controlled evidence does not support using ultrasound for this routine monitoring of RA. However, ultrasound may be useful in assessing disease activity in a narrower subgroup of people with RA; specifically, when clinical examination is inconclusive or is inconsistent with other signs of disease activity (for example, pain or markers of inflammation). Reliable research on the added value of ultrasound in assessing disease activity as part of a monitoring strategy in these subgroups is absent.

If clinical examination is unreliable or uncertain in this subgroup, it will be challenging for healthcare professionals to make a valid clinical assessment and thus apply a treat to target approach and make appropriate management decisions.

In addition, where there is inconsistency between the clinical examination and the disease activity score, it may be unclear if the person has subclinical inflammatory synovitis or more of a widespread pain syndrome, which is not inflammatory. These states require very different treatments, so it is important to define them accurately.

Criteria for selecting high-priority research recommendations:

PICO question	Population: Adults with RA in whom clinical examination is inconclusive or is inconsistent with other signs of disease activity Intervention(s): Treatment adjusted throughultrasound assessment plus usual monitoring assessments Comparison: Treatment adjusted through standard monitoring assessments alone Outcome(s): Disease activity, numbers in remission, numbers with low disease activity, quality of life, function and pain, radiographic progression
Importance to patients or the population	If ultrasound can be used to provide healthcare professionals with additional information on disease activity when standard monitoring assessments are inconclusive, this would enable more informed management decisions to be made. By enabling accurate assessment, ultrasound may facilitate appropriate treatment adjustment to achieve the agreed target and improve prognosis. People with RA may therefore have improvements in clinical status, symptoms and quality of life and avoid receiving inappropriate therapies. Ultrasound is a very simple non-invasive investigation, which is valued by people with RA as it enables them to visualise their disease activity. This in itself may also improve

	outcomes, by encouraging medication adherence and facilitating agreement to treatment escalation where necessary.
Relevance to NICE guidance	Current guidance recommends against using US for routine monitoring of disease activity. This research would aim to identify whether there is an important subset of people with RA who would benefit.
Relevance to the NHS	If ultrasound was found to be clinically and cost effective in assessing disease activity in certain subgroups of people with RA, its use may increase in those groups of people. Although there may be additional training requirements for rheumatologists or other members of the MDT and upfront costs to supply the equipment, if the use of ultrasound for monitoring is found to enable a more appropriate management strategy, this may be cost neutral or even cost saving through better management of RA.
National priorities	N/A
Current evidence base	Randomised controlled evidence in the evidence review reported in chapter I showed no benefit of ultrasound to routine monitoring of disease activity in people with RA. Evidence looking at the association between ultrasound findings and subsequent poor outcomes were of generally low to very low quality and their findings were inconsistent and inconclusive and insufficient to inform a recommendation. Therefore a recommendation was made against the use of ultrasound for routine monitoring for all people with RA. No RCTs were available in the particular subgroups identified in this research question – namely people with RA in whom clinical examination is inconclusive or is inconsistent with other signs of disease activity.
Equality	Ultrasound may be of benefit where synovitis is difficult to assess in case of obesity or extensive deformities.
Study design	Randomised controlled trial comparing treatment adjusted through clinical assessment plus ultrasound, with treatment adjusted through clinical assessment alone. Inclusion criteria would be people with RA in whom clinical examination is inconclusive or is inconsistent with other signs of disease activity, as described above. This could be cluster randomised to aid feasibility by increasing recruitment potential. Trial duration: at least 1 year
Feasibility	The main issue would be to ensure that cross-site agreement on US scoring and technique and therefore this should be considered and pre-specified in the trial protocol.
Other comments	This is an important question appropriate for funding from several potential sources. Ultrasound is currently being used in an ad hoc way with substantial variation in practice around the country. If it is shown to be of added value in particular subgroups of patients, it has the potential to improve the application of the current treat to target strategy in a focussed way.
Importance	• High: the research is essential to inform future updates of key recommendations in the guideline.

1