White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/120676/

Version: Published Version

Article:
Simpson, E.L., Hock, E.S., Stevenson, M.D. et al. (2018) What is the added value of ultrasound joint examination for monitoring synovitis in rheumatoid arthritis and can it be used to guide treatment decisions? A systematic review and cost-effectiveness analysis. Health Technology Assessment, 22 (20). ISSN: 1366-5278

https://doi.org/10.3310/hta22200

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

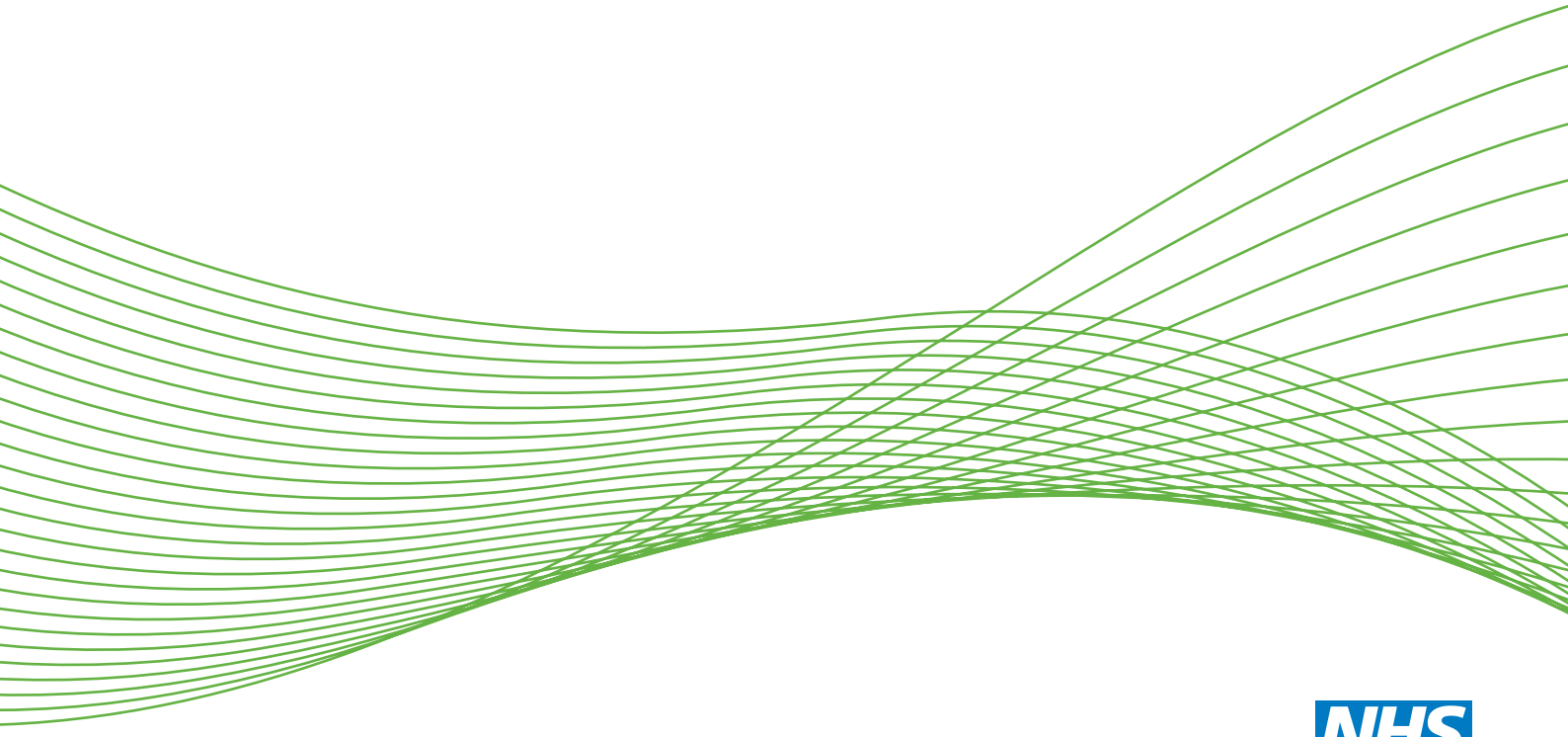
UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

HEALTH TECHNOLOGY ASSESSMENT

VOLUME 22 ISSUE 20 APRIL 2018
ISSN 1366-5278

What is the added value of ultrasound joint examination for monitoring synovitis in rheumatoid arthritis and can it be used to guide treatment decisions? A systematic review and cost-effectiveness analysis

Emma Simpson, Emma Hock, Matt Stevenson, Ruth Wong, Naila Dracup, Allan Wailoo, Philip Conaghan, Cristina Estrach, Christopher Edwards and Richard Wakefield



NHS
National Institute for Health Research

DOI 10.3310/hta22200

What is the added value of ultrasound joint examination for monitoring synovitis in rheumatoid arthritis and can it be used to guide treatment decisions? A systematic review and cost-effectiveness analysis

Emma Simpson,[1]* Emma Hock,[1] Matt Stevenson,[1] Ruth Wong,[1] Naila Dracup,[1] Allan Wailoo,[1] Philip Conaghan,[2,3] Cristina Estrach,[4] Christopher Edwards[5] and Richard Wakefield[2,3]

[1]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[2]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK
[3]National Institute for Health Research (NIHR) Leeds Biomedical Research Centre, Leeds, UK
[4]Aintree University Hospitals NHS Foundation Trust, Liverpool, UK
[5]National Institute for Health Research (NIHR) Wellcome Trust Clinical Research Facility, University of Southampton, Southampton, UK

*Corresponding author

Declared competing interests of authors: Richard Wakefield has provided consulting advice and spoken for General Electric with regard to ultrasound technologies and has also received speaker fees from AbbVie Inc. for ultrasound-related projects. Cristina Estrach has been a member of advisory boards for and/or received speaker fees from AbbVie Inc., Chugai Pharma (UK) Ltd and General Electric Co. Her institution has received educational grants from Pfizer Inc.

Published April 2018
DOI: 10.3310/hta22200

This report should be referenced as follows:

Simpson E, Hock E, Stevenson M, Wong R, Dracup N, Wailoo A, *et al*. What is the added value of ultrasound joint examination for monitoring synovitis in rheumatoid arthritis and can it be used to guide treatment decisions? A systematic review and cost-effectiveness analysis. *Health Technol Assess* 2018;**22**(20).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.236

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 14/16/01. The contractual start date was in January 2015. The draft report began editorial review in April 2016 and was accepted for publication in August 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Health Technology Assessment Editor-in-Chief

Professor Hywel Williams Director, HTA Programme, UK and Foundation Professor and Co-Director of the Centre of Evidence-Based Dermatology, University of Nottingham, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Director of the NIHR Dissemination Centre, University of Southampton, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of members of the NIHR Journals Library Board: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

What is the added value of ultrasound joint examination for monitoring synovitis in rheumatoid arthritis and can it be used to guide treatment decisions? A systematic review and cost-effectiveness analysis

Emma Simpson,[1]* Emma Hock,[1] Matt Stevenson,[1] Ruth Wong,[1] Naila Dracup,[1] Allan Wailoo,[1] Philip Conaghan,[2,3] Cristina Estrach,[4] Christopher Edwards[5] and Richard Wakefield[2,3]

[1]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[2]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK
[3]National Institute for Health Research (NIHR) Leeds Biomedical Research Centre, Leeds, UK
[4]Aintree University Hospitals NHS Foundation Trust, Liverpool, UK
[5]National Institute for Health Research (NIHR) Wellcome Trust Clinical Research Facility, University of Southampton, Southampton, UK

*Corresponding author e.l.simpson@sheffield.ac.uk

Background: Synovitis (inflamed joint synovial lining) in rheumatoid arthritis (RA) can be assessed by clinical examination (CE) or ultrasound (US).

Objective: To investigate the added value of US, compared with CE alone, in RA synovitis in terms of clinical effectiveness and cost-effectiveness.

Data sources: Electronic databases including MEDLINE, EMBASE and the Cochrane databases were searched from inception to October 2015.

Review methods: A systematic review sought RA studies that compared additional US with CE. Heterogeneity of the studies with regard to interventions, comparators and outcomes precluded meta-analyses. Systematic searches for studies of cost-effectiveness and US and treatment-tapering studies (not necessarily including US) were undertaken.

Mathematical model: A model was constructed that estimated, for patients in whom drug tapering was considered, the reduction in costs of disease-modifying anti-rheumatic drugs (DMARDs) and serious infections at which the addition of US had a cost per quality-adjusted life-year (QALY) gained of £20,000 and £30,000. Furthermore, the reduction in the costs of DMARDs at which US becomes cost neutral was also estimated. For patients in whom dose escalation was being considered, the reduction in number of patients escalating treatment and in serious infections at which the addition of US had a cost per QALY gained of £20,000 and £30,000 was estimated. The reduction in number of patients escalating treatment for US to become cost neutral was also estimated.

Results: Fifty-eight studies were included. Two randomised controlled trials compared adding US to a Disease Activity Score (DAS)-based treat-to-target strategy for early RA patients. The addition of power Doppler ultrasound (PDUS) to a Disease Activity Score 28 joints-based treat-to-target strategy in the Targeting Synovitis in Early Rheumatoid Arthritis (TaSER) trial resulted in no significant between-group difference for change in Disease Activity Score 44 joints (DAS44). This study found that significantly more patients in the PDUS group attained DAS44 remission ($p = 0.03$). The Aiming for Remission in Rheumatoid

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Arthritis (ARCTIC) trial found that the addition of PDUS and grey-scale ultrasound (GSUS) to a DAS-based strategy did not produce a significant between-group difference in the primary end point: composite DAS of < 1.6, no swollen joints and no progression in van der Heijde-modified total Sharp score (vdHSS). The ARCTIC trial did find that the erosion score of the vdHS had a significant advantage for the US group ($p = 0.04$). In the TaSER trial there was no significant group difference for erosion. Other studies suggested that PDUS was significantly associated with radiographic progression and that US had added value for wrist and hand joints rather than foot and ankle joints. Heterogeneity between trials made conclusions uncertain. No studies were identified that reported the cost-effectiveness of US in monitoring synovitis. The model estimated that an average reduction of 2.5% in the costs of biological DMARDs would be sufficient to offset the costs of 3-monthly US. The money could not be recouped if oral methotrexate was the only drug used.

Limitations: Heterogeneity of the trials precluded meta-analysis. Therefore, no summary estimates of effect were available. Additional costs and health-related quality of life decrements, relating to a flare following tapering or disease progression, have not been included. The feasibility of increased US monitoring has not been assessed.

Conclusion: Limited evidence suggests that US monitoring of synovitis could provide a cost-effective approach to selecting RA patients for treatment tapering or escalation avoidance. Considerable uncertainty exists for all conclusions. Future research priorities include evaluating US monitoring of RA synovitis in longitudinal clinical studies.

Study registration: This study is registered as PROSPERO CRD42015017216.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of tables

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of figures

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Glossary

Grey-scale, or B-mode, ultrasound Mode of ultrasound that produces a black and white image of a tissue. It demonstrates qualitative differences between tissues and is primarily used to identify synovial hypertrophy and fluid within joints.

Power Doppler ultrasound Mode of ultrasound that detects blood flow. Flow is designated by a colour (often orange or red), which is overlaid onto the corresponding grey-scale image below. There are different types of Doppler but power Doppler is thought to be most sensitive to the relatively low blood flow seen in joints.

Sensitivity Proportion of true-positive results, a measure of the accuracy of a diagnostic test.

Spearman's rho (ρ) Spearman's rank correlation coefficient, a measure of the association between two variables.

Specificity Proportion of true-negative results, a measure of the accuracy of a diagnostic test.

Standardised response mean The ratio of the mean changes to the standard deviation of the changes, a measure of responsiveness to change.

Synovitis Swelling of the synovial lining of the joints.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of abbreviations

ABT	abatacept
ACR	American College of Rheumatology
ADA	adalimumab
ARCTIC	Aiming for Remission in Rheumatoid Arthritis
bDMARD	biological disease-modifying anti-rheumatic drug
BSR	British Society for Rheumatology
CDAI	Clinical Disease Activity Index
cDMARD	conventional disease-modifying anti-rheumatic drug
CDSR	Cochrane Database of Systematic Reviews
CE	clinical examination
CENTRAL	Cochrane Central Register of Controlled Trials
CI	confidence interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CRP	C-reactive protein
CT	computerised axial tomography
CTZ	certolizumab pegol
CXCL13	chemokine (C-X-C motif) ligand 13
DARE	Database of Abstracts of Reviews of Effects
DAS	Disease Activity Score
DAS28	Disease Activity Score 28 joints
DAS28-CRP	Disease Activity Score 28 joints using C-reactive protein
DAS28-ESR	Disease Activity Score 28 joints using erythrocyte sedimentation rate
DAS44	Disease Activity Score 44 joints
DMARD	disease-modifying anti-rheumatic drug
ESR	erythrocyte sedimentation rate
ETN	etanercept
EULAR	European League Against Rheumatism
FN	false negative
FP	false positive
GATE	Generic Appraisal Tool for Epidemiology
GOL	golimumab
GSUS	grey-scale ultrasound
HAQ	Health Assessment Questionnaire
HAQ-DI	Health Assessment Questionnaire Disability Index
HCQ	hydroxychloroquine
HTA	Health Technology Assessment
ICER	incremental cost-effectiveness ratio
IFX	infliximab
MCP	metacarpophalangeal
MRI	magnetic resonance imaging
MTP	metatarsophalangeal
mTSS	modified total Sharp score
MTX	methotrexate
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NPV	negative predictive value
NRAS	National Rheumatoid Arthritis Society
NSAID	non-steroidal anti-inflammatory drug
OMERACT	Outcome Measures in Rheumatoid Arthritis Clinical Trials

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

OR	odds ratio
PDUS	power Doppler ultrasound
PET	positron emission tomography
PIP	proximal interphalangeal
PPV	positive predictive value
QALY	quality-adjusted life-year
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
QUIPS	Quality in Prognosis Studies
RA	rheumatoid arthritis
RAMRIS	rheumatoid arthritis magnetic resonance imaging scoring system
RAQoL	Rheumatoid Arthritis Quality of Life
RCT	randomised controlled trial
RTX	rituximab
ScHARR	School of Health and Related Research
SDAI	Simplified Disease Activity Index
SJC	swollen joint count
SRM	standardised response mean
SSZ	sulfasalazine
TA	technology appraisal
TaSER	Targeting Synovitis in Early Rheumatoid Arthritis
TCZ	tocilizumab
TJC	tender joint count
TN	true negative
TNFi	tumour necrosis factor inhibitor
TP	true positive
TSS	total Sharp score
TTT	treat to target
TURA	Targeted Ultrasound in Rheumatoid Arthritis
US	ultrasound
US7	ultrasound 7 score
USDCF	Ultrasound Doppler colour fraction
VAS	visual analogue scale
vdHSS	van der Heijde-modified total Sharp score

Plain English summary

The aims of this study were to investigate the value of ultrasound (US) in addition to clinical examination (CE) for monitoring synovitis (swelling of joint synovial lining) in rheumatoid arthritis (RA) and whether or not US could be used to make treatment decisions. A systematic literature review identified 58 studies providing relevant evidence. Studies indicated that US-detected synovitis was associated with later development of structural progression assessed by radiography, a sign that RA is causing permanent joint damage. Studies suggested that US was better than CE at predicting response to treatment and tapering (dose reduction) or discontinuation of a particular treatment. US was used to make treatment decisions and could increase clinician and patient confidence in those decisions. However, studies varied in the types of US and CE that were used and how outcomes were measured, making it difficult to draw firm conclusions about the overall usefulness of adding US to CE.

No studies on the cost-effectiveness of US for monitoring synovitis were identified, but studies assessing the impact of reducing treatment doses were identified. These reported that some patients who had achieved good control of their disease could have their treatment reduced without harmful effects. It may also be possible to withdraw treatment for some patients although this is rare in established RA. Reduced dosages can save large sums of money in terms of drug costs. It was estimated that a 2.5% reduction in dose of biological treatment would more than pay for the use of US every 3 months. The reduction in dose needed to cover the costs of US rose when less expensive treatments were used but did not become larger than 6%, assuming that multiple cheaper treatments were used. It appears likely that using US to monitor synovitis could potentially represent value for money. However, more evidence is needed to reduce the uncertainty in the current findings.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Scientific summary

Background

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterised by progressive, irreversible joint damage, impaired joint function, pain and tenderness caused by swelling of the synovial lining of joints (synovitis) and is manifested by increasing disability and reduced quality of life. RA is associated with substantial costs both directly (associated with drug acquisition and hospitalisation) and indirectly, because of reduced productivity. Synovitis is assessed by clinical examination (CE) of the joints. Synovitis can also be assessed using imaging technologies including magnetic resonance imaging (MRI) and ultrasound (US); this may detect synovitis that is not detected by CE (subclinical synovitis) and may also distinguish between synovitis and other pathologies more readily than CE alone. US can therefore aid key decision-making with regard to therapy changes, leading to escalation or tapering of therapy. This is important for a disease area in which modern therapies are expensive and all therapies are associated with side effects, especially infections.

Objective

This report aimed to address the question: 'What is the added value of US joint examination for monitoring synovitis in RA and can it be used to guide treatment decisions?'

Data sources

The following electronic databases were searched: MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid) (1946 to October 2015), EMBASE (via Ovid) (1974 to October 2015), Cochrane Database of Systematic Reviews (CDSR) (1996 to October 2015), Cochrane Central Register of Controlled Trials (CENTRAL) (1898 to October 2015), Health Technology Assessment (HTA) database (1989 to October 2015), Database of Abstracts of Reviews of Effects (DARE) (1946–2014), NHS Economic Evaluation Database (NHS EED) (1968–2014), Science Citation Index Expanded (1900 to October 2015), Science Citation Index and Conference Proceedings Index (1900 to October 2015), ClinicalTrials.gov (October 2015), European League Against Rheumatism Abstract Archive (via Web of Science) (October 2015), American College of Rheumatology and Association of Rheumatology Health Professionals (via Web of Science) (October 2015) and Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) conference proceedings (via Web of Science) (October 2015).

Review methods

A systematic review of US was conducted. Studies were sought that compared grey-scale US (GSUS) or power Doppler US (PDUS) with CE, the use of inflammatory biomarkers or disease activity scoring tools. The patient group considered was adults with RA. Outcomes included diagnostic (detection of synovitis, responsiveness to change), prognostic (association with progression or other disease outcomes) or treatment-related (response to treatment, treatment tapering, influence on decisions) measures. Diagnostic studies, prognostic studies and studies investigating the prediction of response to treatment, or treatment tapering, or the influence of US on treatment decisions were included. Although there are other methods for detecting synovitis, the decision was taken to limit the intervention in the review to US. Study selection was carried out by two reviewers, with disagreements resolved by discussion. Included studies were quality assessed and data extracted by one reviewer and checked by another reviewer. Data were tabulated and discussed. Study heterogeneity precluded meta-analyses.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

A review of the cost-effectiveness of the use of US to monitor synovitis and a systematic search for the outcomes associated with tapering of RA treatment, irrespective of whether or not US was used, were undertaken. Study selection was carried out by two reviewers: any study deemed relevant by at least one reviewer was retrieved. Included studies were summarised by one reviewer. The tapering search was supplemented by checking the reference lists of included studies, searching for subsequent publications related to any abstracts identified and retrieving papers known to our clinical advisors. Relevant studies were summarised separately for cost-effectiveness and tapering. A survey publicised to UK rheumatology units was undertaken to investigate whether or not US is being used to monitor synovitis and guide treatment decisions in RA.

Mathematical model

The modelling undertaken was purposefully simplistic so that the key interactions between monitoring synovitis with US and decisions influencing treatment could be examined explicitly. The simple model estimated for patients in whom the clinician was contemplating reducing the treatment dose included (1) the reduction in treatment costs and therapy modification leading to serious infection avoidance that would be required for the addition of US for monitoring synovitis to have a cost per quality-adjusted life-year (QALY) gained of £20,000 and £30,000 and (2) the reduction in treatment costs required for the addition of US for monitoring synovitis to become cost saving. For patients in whom the clinician was thinking of increasing the treatment dose, two analyses were undertaken: (1) the reduction in the number of patients not progressing to more intensive treatment therapy or avoiding a serious infection through the use of US needed to achieve cost per QALY gained values of £20,000 and £30,000 and (2) the reduction in the number of patients escalating treatment needed for US to become cost saving.

Results

In the systematic review, 2724 records were identified from the electronic databases and an additional 26 records were identified from bibliography searching. Following title and abstract sifting, 154 articles were assessed for eligibility, of which 63 full-text papers were excluded. In total, 75 articles describing 58 studies were included; additionally, one study identified by the search as ongoing was published prior to publication of this report. A further 16 articles were retained for bibliography checking. Twenty-six studies provided prognostic and/or treatment data and 32 studies provided diagnostic data only.

Two randomised controlled trials (RCTs) of the treatment strategy did not find significant benefits in terms of the primary outcome of adding US to a Disease Activity Score (DAS)-based treat-to-target strategy for early RA patients. The addition of PDUS to a Disease Activity Score 28 joints-based treat-to-target strategy in the Targeting Synovitis in Early Rheumatoid Arthritis (TaSER) RCT resulted in no significant between-group difference in change from baseline in Disease Activity Score 44 joints (DAS44) (mean change: intervention –2.69, control –2.58). This study found that the addition of PDUS to the treatment strategy led to significantly more patients attaining DAS44 remission (66%) than for the DAS-alone strategy (43%) ($p = 0.03$). The Aiming for Remission in Rheumatoid Arthritis (ARCTIC) RCT found that the addition of PDUS and GSUS to a DAS-based strategy did not produce a significant between-group difference in the primary end point [which consisted of a composite of DAS of < 1.6, no swollen joints at 16, 20 and 24 months and no progression in van der Heijde-modified total Sharp score (vdHSS) between 16 and 24 months], with values of 22.0% and 18.8%, respectively. The ARCTIC trial did find that the change in erosion score of the vdHS had a significant advantage for the US group over the DAS group (changes of 0.5 and 1.0 for the US and control groups, respectively; $p = 0.04$). Erosion in the TaSER trial, as measured by change in the erosion score of the rheumatoid arthritis magnetic resonance imaging scoring system (RAMRIS), was not significantly different between the groups (changes of 0.5 and 1.0 for the US and control groups, respectively). These studies did not however explore the value of US when added only in cases of clinical uncertainty, for example, when there was discrepancy between DAS and clinical evaluation.

The majority of prospective cohort studies investigating radiographic progression reported that US at baseline, either GSUS or PDUS, was significantly correlated with progression at follow-up ($p = 0.05$ to $p < 0.001$). Radiographic progression in most, but not all, cases was measured with a modification of the total Sharp score/vdHSS. PDUS was significantly associated with radiographic progression in 10 studies in which it was measured. Associations were reported as odds ratios (ORs) in two studies and were 12.21 [95% confidence interval (CI) 3.34 to 44.73; $p < 0.001$] and 1.80 (95% CI 1.20 to 2.71; $p = 0.005$). Associations reported as correlation coefficients ranged from $r = 0.099$ to $r = 0.77$ ($p = 0.05$ to $p < 0.001$). Significance levels from Mann–Whitney *U*-test results reported were $p < 0.001$ and $p = 0.0011$. GSUS was significantly associated with radiographic progression in six studies but not in three studies. Significant associations reported were Mann–Whitney *U*-test $p = 0.027$; $r = 0.140$ to $r = 0.61$ ($p < 0.001$); and ORs of 2.08 (95% CI 1.39 to 3.11) and 2.15 (95% CI 1.23 to 3.75) ($p = 0.01$). The difference between studies reporting significant associations and studies reporting non-significant associations could not be explained by study quality, joints assessed or how the end point was measured.

Other outcomes reported were heterogeneous, making it difficult to draw conclusions. PDUS was significantly correlated with the proportion of patients experiencing disease flare at follow-up ($p = 0.014$), whereas this was not significant for GSUS. US could significantly predict treatment persistence ($p = 0.02$) and US predicted treatment tapering or discontinuation failure significantly in two out of three studies ($p = 0.005$, $p < 0.0005$, $p = 0.06$) whereas clinical measures alone did not. The additional use of US modified treatment decisions (in 23–88% of cases in UK studies) and significantly increased ($p < 0.001$ to $p < 0.0005$) clinician confidence in treatment decisions.

The review of cost-effectiveness studies identified five articles, although none was directly relevant to the decision problem. Nineteen papers were identified from the tapering search, which, when supplemented by checking references of identified articles, by searching for subsequent articles related to identified abstracts and by articles known to our clinical advisors, resulted in 39 relevant papers being included. Given that evidence showed that some patients who had achieved a low level of disease activity could have their treatment tapered, with no or little short-term harm to them, it was deemed appropriate to model strategies in which a clinician was contemplating a dose reduction. The survey conducted yielded only 31 responses, 27 of which stated that US was used for treatment decisions. The small sample size means that the results can not be generalised across the UK.

For patients who have been stable on biological disease-modifying anti-rheumatic drugs (bDMARDs) and in whom the clinician is contemplating reducing the dose of bDMARDs, the model estimated that an average reduction of 2.5% in the costs of bDMARDs was sufficient to cover the costs of performing US every 3 months. Similarly, if 2.5% of patients do not have their treatment escalated to bDMARDs, it was estimated that the use of US to monitor synovitis would be cost neutral. If only conventional disease-modifying anti-rheumatic drugs (DMARDs) were considered in the current or planned treatment regimen, the money spent on US monitoring could not be recouped.

Limitations

Few RCTs were available and so lower-quality study designs were included in the review. The heterogeneity of the studies identified precluded meta-analysis. Therefore, no summary estimates of effect were available, which is a limitation of the review. There is no gold standard/reference standard for the detection of synovitis, although it has been suggested that MRI may be used as a gold standard/reference standard; it detects similar levels of inflammation to US. The systematic search for DMARD tapering was not overly sensitive as many articles were identified by bibliography and citation searching and contact with clinical advisors (rather than by database searching). However, the summarised articles provide a broad overview of the literature base and it is likely that any papers omitted would not have an adverse impact on the findings of this report.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Limitations in the modelling include the exclusion of biosimilar bDMARDs, which are likely to reduce the cost associated with bDMARD treatment compared with the base case. However, a sensitivity analyses was conducted to explore the impact of price reductions on the threshold levels.

A key limitation within the modelling was the lack of robust data relating to key parameters within the decision problem. As such, threshold analyses have been undertaken to provide indicative levels of drug dose reduction and avoidance of serious infections required to make the use of US cost neutral.

Conclusion

Limited evidence was available and therefore cost-effectiveness analysis was limited to threshold analysis. Given the proportions of patients who could potentially taper treatment, or remain on stable therapy without escalation, the use of US to monitor synovitis could potentially be a cost-effective approach if it provides clinicians with more confidence in reducing the drug burden. However, there is considerable uncertainty in this conclusion.

The most important future research studies would be longitudinal studies evaluating the role of US in the management of synovitis in RA.

Study registration

This study is registered as PROSPERO CRD42015017216.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Background

Description of the health problem

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterised by progressive, irreversible joint damage, impaired joint function, pain and tenderness caused by swelling of the synovial lining of the joints (synovitis) and is manifested with increasing disability and reduced quality of life.[1] The primary symptoms are pain, morning stiffness, swelling, tenderness, loss of movement, fatigue and redness of the peripheral joints.[2,3] RA is associated with substantial costs both directly (associated with drug acquisition and hospitalisation) and indirectly, because of reduced productivity.[4] RA has long been reported as being associated with increased mortality,[5,6] particularly as a result of cardiovascular events.[7]

Diagnosis of rheumatoid arthritis

The initial classification criteria for RA were produced in 1987 by the American College of Rheumatology (ACR).[8] National Institute for Health and Care Excellence (NICE) clinical guideline 79[9] provides a summary of the ACR criteria, with patients needing to have at least four of the seven criteria to be given a diagnosis of RA: morning stiffness lasting at least 1 hour, swelling in three or more joints, swelling in the hand joints, symmetric joint swelling, erosions or decalcification on radiography of the hand, rheumatoid nodules and abnormal serum rheumatoid factor. For the first four criteria these must have been present for at least a period of 6 weeks. However, in the clinical guideline[9] the guideline development group preferred a clinical diagnosis of RA rather than the ACR criteria to permit early treatment, as consistent with European League Against Rheumatism (EULAR) recommendations.[10]

In 2010, the ACR and EULAR jointly published RA classification criteria, which focused on features present at earlier stages of the disease that are associated with persistent and/or erosive disease rather than defining the disease by its late-stage features.[11] The classification criteria allocate scores to the characteristics of joint involvement, serology, acute-phase reactants and duration of symptoms to produce a total score between 0 and 10, with those scoring ≥ 6 and with obvious clinical synovitis being defined as having 'definite RA' in the absence of an alternative diagnosis that better explains the synovitis.[11] The growing recognition of the accuracy of modern imaging [such as magnetic resonance imaging (MRI) and ultrasound (US)] over clinical examination (CE) in detecting synovitis was highlighted in these new classification criteria, which allow the extent of joint involvement to be determined by imaging.[11]

Epidemiology

There are an estimated 400,000 people in England and Wales with RA,[12] with approximately 10,000 incident cases per year.[13] The disease is more prevalent in females (1.16%) than in males (0.44%),[13] with the majority of cases being diagnosed when patients are aged between 40 and 80 years[14] and a peak incidence in those in their 70s.[13]

Measurement of disease activity and damage progression

Synovitis can be detected by CE. Monitoring may involve taking swollen joint counts (SJCs) or tender joint counts (TJCs). Biomarkers may be used to detect evidence of systemic inflammation, for example anaemia in chronic disease or elevated levels of C-reactive protein (CRP) or an elevated erythrocyte sedimentation rate (ESR).[9,15] Common measures of disease activity, disability and response to treatment are presented in *Table 1*. ACR response[19] and EULAR response[20] measures have dominated the measurement of improvement in RA symptoms. The ACR response has been widely adopted in randomised controlled trials (RCTs) although studies have shown that the value can vary between trials because of the timing of the response.[21] The EULAR response criteria and the ACR20 improvement criteria (see *Table 1*) were found to have reasonable agreement in the same set of clinical trials, although van Gestel *et al.*[22] stated that the EULAR response criteria showed better construct and discriminant validity than the ACR20 criteria. The Disease Activity Score 28 joints

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 1 Measurement of disease activity and damage progression

Measure	Description
Health Assessment Questionnaire (HAQ)[16]	A patient-completed disability assessment with established reliability and validity. Scores range from 0 to 3, with higher scores indicating greater disability. The HAQ scale is a discrete scale with step values of 0.125, resulting in 25 points on the scale
Visual analogue scale (VAS)	A global assessment of the extent of disease, which may be assessed by the patient or the physician. This method may also be used for patient-reported pain. The VAS usually consists of a 100-mm line that represents the total range of the disease extent/pain levels experienced. Patients/physicians mark a point on the line
DAS28-ESR[17]	This assesses 28 joints (shoulder, knee, elbow, wrist, MCP joints one to five, PIP joints one to five, bilaterally) in terms of swelling (SW28) and of tenderness to the touch (TEN28). It also incorporates measures of the ESR and a subjective assessment (SA) on a scale of 0–100 made by the patient regarding disease activity in the previous week. The DAS28-ESR score is calculated using the following equation: $0.56 \times TEN28^{0.5} + 28 \times SW28^{0.5} + 0.70 \times \ln(ESR) + 0.014 \times SA$. A DAS28-ESR of ≤ 3.2 indicates inactive disease; a DAS28-ESR of > 3.2 and ≤ 5.1 indicates moderate disease; a DAS28-ESR of > 5.1 indicates very active disease
DAS28-CRP[17]	Similar to the DAS28-ESR but uses the CRP value instead of the ESR. The DAS28-CRP is calculated using the following equation: $[0.56 \times \text{sqrt}(TEN28) + 0.28 \times \text{sqrt}(SW28) + 0.36 \times \ln(CRP + 1)] \times 1.10 + 1.15$
Clinical Disease Activity Index (CDAI)[18]	A composite index based on SJCs (0–28), TJCs (0–28), CRP, patient global assessment of disease activity VAS (0–10 cm) and physician global assessment of disease activity VAS (0–10 cm)
Simplified Disease Activity Index (SDAI)[18]	A simplified version of the CDAI based on SJCs (0–28), TJCs (0–28), patient global assessment of disease activity VAS (0–10 cm) and physician global assessment of disease activity VAS (0–10 cm)
ACR response[19]	An ACR20 response requires a 20% improvement in TJCs; a 20% improvement in SJCs; and a 20% improvement in at least three of the following five 'core set items': physician global assessment, patient global assessment, patient pain, self-reported disability (using a validated instrument) and ESR or CRP. ACR50 and ACR70 responses require 50% and 70% improvements, respectively
EULAR response[20]	Participants are classified as good, moderate or non-responders based on the individual change in DAS28 and the DAS28 reached.[20] The relationship between change in DAS28 and DAS28 reached and EULAR response is shown in *Table 2*

DAS28-CRP, Disease Activity Score 28 joints using C-reactive protein; DAS28-ESR, Disease Activity Score 28 joints using erythrocyte sedimentation rate; MCP, metacarpophalangeal; PIP, proximal interphalangeal; sqrt, square root.

(DAS28; see *Table 1*) can be used to classify both the disease activity of the patient and the level of improvement estimated. The EULAR response has been reported less frequently in RCTs than the ACR response, although the EULAR criteria are much more closely aligned to the treatment continuation rules stipulated by NICE,[9] which require a DAS28 improvement of > 1.2 to continue treatment.[22]

The relationship between change in DAS28 and DAS28 reached and EULAR response is shown in *Table 2*. Depending on the initial DAS28 of the patient, this change in DAS28 would equate to either a good or moderate EULAR response, as shown in the second column of *Table 2*.

Both the ACR criteria and the DAS28-based EULAR criteria rely on subjective measurements and can be influenced by patients' pain threshold and comorbidities.[24] The DAS28 has been criticised as patients achieving DAS remission can still have synovitis as evidenced by imaging.[25,26] Alternatively, chronic pain experienced by patients or deformity and residual fibrous tissue leading to the detection of swelling by manual examination of joints can raise the DAS, thus overestimating disease activity.[24] If DAS measurement is based on CRP levels, it can be influenced by therapies; for example, tocilizumab (TCZ; RoActemra®, Roche Products Ltd, Welwyn Garden City, UK) has a direct effect on CRP through interleukin-6 and can therefore artificially lower the DAS28.[27] Unlike these subjective measurements, imaging could provide an objective measurement of synovitis.

TABLE 2 Determining the EULAR response based on DAS28[23]

	Improvement in DAS28[23]		
DAS28 at end point	> 1.2	> 0.6 and ≤ 1.2	≤ 0.6
≤ 3.2	Good	Moderate	None
> 3.2 and ≤ 5.1	Moderate	Moderate	None
> 5.1	Moderate	None	None

Notes
The shaded cells indicate where patients continue treatment based on current NICE technology appraisal guidance.[9]
This table contains information from Stevenson *et al.*,[23] licensed under the Non-Commercial Government Licence v2.0.

The role of imaging

Imaging techniques used in the diagnosis and monitoring of RA include US, conventional radiography, MRI and computerised axial tomography (CT). Initial diagnosis of RA, or differential diagnosis of forms of arthritis, may be aided by imaging. There is a EULAR recommendation that US, MRI and conventional radiography can improve the certainty of a RA diagnosis if there is diagnostic doubt.[28]

In terms of detecting inflammation, there is evidence to demonstrate that both US and MRI detect more cases of synovitis than CE.[9,28] The benefit of US as an addition to CE will be influenced by the ability of this subclinical synovitis to predict disease progression. Prognostic studies of US-detected synovitis were sought in this report (see *Chapter 3*). Because of their three-dimensional acquisition, both US and MRI are also more sensitive than conventional radiography at detecting erosion damage and early signs of erosion.[9] MRI can detect bone oedema, which may be a predictor of radiographic progression.[28]

Although MRI may be more sensitive than US for detecting erosions,[29] US and MRI have comparable abilities to detect synovitis.[28–30] Arthritis Research UK[30] suggests that US has an advantage over other imaging techniques as it can be used immediately after CE to assess symptomatic areas or clinical abnormalities and this has the added advantage of improving clinical skills.[30]

There is no conclusive gold standard (sometimes referred to as a reference standard) for assessing synovitis. When compiling their 2009 guidelines, NICE[9] recognised that many clinicians had limited access to US and so it considered clinician examination as the gold standard for the detection of synovitis in practice. It is possible that US has become more widespread since then. To address this issue, a survey was conducted to determine the usage of US in UK clinical practice (see *Appendix 1*).

Conventional radiography is commonly used to measure disease progression.[9] In clinical trials this may be carried out using the total Sharp score (TSS), the van der Heijde-modified total Sharp score (vdHSS)[31] or the modified Genant *et al.*[32] scoring method, which measure erosions and joint space narrowing, although systematic scoring of radiographs in routine clinical care is very uncommon.

Treatment strategies based on a therapeutic outcome target (known as 'treat to target'; TTT) are becoming more widely used within rheumatology.[33–35] Typically, targets are either remission or low disease activity (when remission is not attainable). These are usually clinically defined [e.g. DAS of < 2.6, Simplified Disease Activity Index (SDAI) score of ≤ 3.3].[34,36] However, as a previous review has reported that US has been shown to be superior to CE for detecting synovitis, providing that this is linked to later outcomes,[28] US-defined remission may be a more appropriate target for TTT strategies. This previous review thoroughly examined the use of US to detect synovitis;[28] however, it was published prior to the publication of the treatment studies, including RCTs, in this report. Research is starting to examine the effectiveness of US compared with clinical treatment targets within TTT strategies in RA.[36,37]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Significance for the NHS

Because previous NICE technology appraisals have recommended a number of biological disease-modifying anti-rheumatic drugs (bDMARDs) (see *Current service provision*), with a potential sequence of three bDMARDs, there has been a considerable increase in expenditure on RA interventions as bDMARDs are markedly more expensive than conventional disease-modifying anti-rheumatic drugs (cDMARDs). Any imaging technology that could inform decisions on when to start, or to taper or discontinue, treatments has the potential to benefit the NHS. The frequency of monitoring would influence cost-effectiveness.

Further detailed information on RA can be found in NICE clinical guideline 79.[9] Additional information can also be located in the British Society for Rheumatology (BSR) guidelines.[38]

Current service provision

Traditionally, patients have been treated with cDMARDs, which include methotrexate (MTX), sulfasalazine (SSZ), hydroxychloroquine (HCQ), leflunomide, ciclosporin and gold injections as well as corticosteroids, analgesics and non-steroidal anti-inflammatory drugs (NSAIDs). However, more recently, a group of drugs has been developed consisting of monoclonal antibodies and soluble receptors that specifically modify the disease process by blocking key protein messenger molecules (such as cytokines) or cells (such as B-lymphocytes).[9] Such drugs have been labelled as bDMARDs.

A simplified model of the RA clinical pathway has been provided by NICE,[39] with the use of US falling into the monitoring and review box. For the purposes of this report, modelling assumes that US would be used only at points in the pathway where a change of treatment is being considered, specifically for those patients for whom a clinician is considering tapering treatment to assess the current level of synovitis and for those patients for whom a clinician is contemplating escalating treatment to rule out non-synovitis-related reasons.

Clinical guidelines

For people with newly diagnosed RA, NICE clinical guideline 79[9] recommends a combination of cDMARDs [including MTX and at least one other disease-modifying anti-rheumatic drug (DMARD) plus short-term glucocorticoids] as first-line treatment, ideally beginning within 3 months of the onset of persistent symptoms. When combination therapies are not appropriate (e.g. when there are comorbidities or the patient is pregnant), DMARD monotherapy is recommended. When DMARD monotherapy is used, emphasis should be on increasing the dose quickly to obtain the best disease control. For the purposes of this report, the term 'intensive DMARDs' has been used to denote treatment with multiple cDMARDs simultaneously.

Current National Institute for Health and Care Excellence technology appraisal guidance

National Institute for Health and Care Excellence guidance [technology appraisal (TA) 375][40] recommends the use of the tumour necrosis factor inhibitors (TNFis) etanercept (ETN; Enbrel®, Pfizer Ltd, Sandwich, UK), infliximab (IFX; Remicade®, Merck Sharp & Dohme Ltd, Hoddesdon, UK), adalimumab (ADA; Humira®, AbbVie Ltd, maidenhead, UK), certolizumab pegol (CTZ; Cimzia®, UCB Pharma, Slough, UK) and golimumab (GOL; Simponi®, Merck Sharp & Dohme Ltd), in combination with MTX, in people with RA after the failure of two cDMARDs, including MTX, and who have a DAS28 of > 5.1 and when bDMARDs have not been tried previously. TA375 also recommends the use of TCZ and abatacept (ABT; Orencia®, Bristol-Myers Squibb, Uxbridge, UK) as alternatives to TNFis in the same circumstances.

Technology appraisal 375 did not recommend the use of bDMARDs in patients with a DAS of ≤ 5.1 nor in patients in whom cDMARDs have failed.[40]

The National Institute for Health and Care Excellence has also issued guidance on the treatment of RA after the failure of a TNFi (TA195,[41] TA225[42] and TA247[43]).

National Institute for Health and Care Excellence clinical guideline 79[9] recommends the use of intensive cDMARDs – two cDMARDs used in combination, rather than two cDMARDs used sequentially – although this latter option is acceptable.

A simplified summary of the typical pathway recommended by NICE would be the use of intensive cDMARDs followed by a bDMARD, followed by rituximab (RTX; MabTheraR, Roche Products Ltd, Welwyn Garden City, UK) plus MTX and then TCZ before returning to cDMARDs. If RTX or MTX is contraindicated then ADA, ETN, IFX or ABT in combination with MTX or ADA or ETN monotherapy can be used instead (TA195[41]), as can TCZ in combination with MTX (TA247[43]) and GOL in combination with MTX (TA225[42]). A NICE single technology appraisal recommended the use of CTZ after an inadequate response to a TNFi for patients with a DAS28 of > 5.1 for whom RTX is contraindicated or not tolerated.[44]

National Institute for Health and Care Excellence criteria for continuing treatment

Each of the NICE TAs state that for patients to continue treatment with a bDMARD there must have been an improvement in DAS28 of at least 1.2 points at 6 months. If this criterion has not been met, treatment should be stopped and the next intervention in the sequence initiated.

Current service cost

The decision problem compares the use of US for the monitoring of synovitis in addition to CE compared with CE alone, assuming that there are US facilities within the hospital. As such, no attempt has been made to estimate the costs of providing a rheumatology service without US monitoring, as it is presumed that these costs, such as maintenance costs, consumable costs and receptionist costs, will be the same regardless of whether or not US monitoring is undertaken for RA patients. No analyses have been undertaken assuming that the machinery required to perform US would be bought primarily for the use of RA patients, although it is expected that in such circumstances the costs per US scan would be markedly higher.

Description of the technology under assessment

In US, reflected pulses of high-frequency sound are used to assess soft tissue, cartilage and bone surfaces.[30]

Grey-scale ultrasound (GSUS) (also known as B-mode ultrasound) displays different intensities of echoes, denoting different densities of tissue, in black, white and shades of grey.[30] GSUS can measure synovial hypertrophy and effusion.[45] Doppler US is based on the principle that sound waves increase or decrease in frequency when objects (such as blood cells) move towards or away from the transducer, respectively.[30] Thus, power Doppler ultrasound (PDUS) can be used to detect the volume of blood flow.[30]

Musculoskeletal US can be used for evaluating joint and soft tissue pathology.[30] Within the field of rheumatology, US can be used for the initial diagnosis of arthritis and for US-guided injections or aspirations (which fall outside the scope of this report).[30] It can also be used to assess the extent of inflammatory disease and determine the response to therapy.[30] A significant advantage of US over other imaging technologies such as MRI and CT is the ability to focus on the area of symptoms or clinical abnormality with the US probe immediately after CE.[30]

The ACR has published guidelines on the use of US in rheumatology.[46] This included several recommendations across a range of disorders and various stages of the clinical pathway, including that it is reasonable to use US to evaluate inflammatory disease activity.[46]

The EULAR recommendations on imaging for RA, published in 2013, covered several imaging technologies and pathologies across stages of the clinical pathway.[28] Recommendations included that US be used to detect and monitor inflammation and that imaging be used to predict the response to therapy.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Ultrasound is useful for distinguishing inflammatory from non-inflammatory joint disease. Effusion may indicate synovial inflammation; however, abnormally thickened hypoechoic intra-articular tissue may indicate synovitis in the absence of effusion.[30] PDUS can be used to assess synovitis as it can detect minimal increases in perfusion in the synovium.[30] US can detect subclinical synovitis.[28,30] Subclinical synovitis may predict structural deterioration in RA patients; thus, prognostic studies were sought (see *Chapter 3*) as progression can occur in RA patients in remission.[47,48] The benefit of US as an addition to CE will be influenced by the ability of this subclinical synovitis to predict disease progression.

Interpreting US scans is a skill. In the past, there have been concerns about inter-rater reliability; however, a review published in 2010[49] found that both intra-rater and inter-rater reliability for still-image interpretation was high for both GSUS and PDUS. Limited evidence is available for assessing image acquisition reliability.[50] The Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) US task force has addressed many issues involved in measuring RA and has looked specifically at the variability of US.[30,51] In the past, GSUS was used to assess the presence or absence of synovitis, but this was followed by the creation of semiquantitative scoring systems (*Table 3*).[45] PDUS used to detect vascularity may be quantified[30] (see *Table 3*). Concerns had also been voiced regarding intermachine variability, but with current high-end systems this is now less of a concern.[30,45] D'Agostino *et al.*,[60] as reported by Arthritis Research UK,[30] showed that there is good reliability between experts using different types of US machine. There may be some variability in synovitis throughout the day; response to cDMARDs may be detected within 6–8 weeks and response to bDMARDS may be seen as early as 1 week.[61,62] The optimal frequency of testing will depend on the frequency of treatment decisions. It is envisaged that the number of US tests conducted per year would be aligned to the frequency at which clinical decisions regarding changing the treatment for a patient would be undertaken.

At the OMERACT 7 meeting in 2004 an international panel of experts agreed on the first definitions of sonographic pathology.[51] This meeting achieved consensus in US definitions of joint effusion, bursal effusion and synovitis. Studies following on from that meeting, undertaken within the EULAR/OMERACT network, were reported at the OMERACT 8 meeting in 2006.[63] The international panel of experts

TABLE 3 Ultrasound semiquantitative scoring systems

Reference	Scoring system
Szkudlarek *et al.* 2003[52]	Joint effusion (compressible anechoic intracapsular area): • 0 = no effusion • 1 = minimal amount of joint effusion • 2 = moderate amount of joint effusion (without distension of the joint capsule) • 3 = extensive amount of joint effusion (with distension of the joint capsule) Synovitis (non-compressible hypoechoic intracapsular area): • 0 = no synovial thickening • 1 = minimal synovial thickening (filling the angle between the periarticular bones, without bulging over the line linking the tops of the bones) • 2 = synovial thickening bulging over the line linking the tops of the periarticular bones but without extension along the bone diaphysis • 3 = synovial thickening bulging over the line linking the tops of the periarticular bones and with extension to at least one of the bone diaphyses Power Doppler (flow signal in the synovium): • 0 = no flow in the synovium • 1 = single vessel signals • 2 = confluent vessel signals in less than half of the area of the synovium • 3 = vessel signals in more than half of the area of the synovium

TABLE 3 Ultrasound semiquantitative scoring systems (*continued*)

Reference	Scoring system
Scheel *et al.* 2005[53]	Synovitis and effusion included in a combined scoring method (the larger the anechoic structure or extent of synovial hypertrophy seen on US images, the higher the assigned score): • 0 = no effusion/hypertrophy (no anechoic, hypoechoic or hyperechoic structure visible) • 1 = minimal effusion/hypertrophy • 2 = moderate effusion/hypertrophy • 3 = extensive effusion/hypertrophy
Naredo *et al.* 2005,[54] Naredo *et al.* 2007,[55] Naredo *et al.* 2008[56]	Subjective grading of joint effusion and synovitis from 0 to 3: • 0 = absence • 1 = mild • 2 = moderate • 3 = marked Subjective grading of the intra-articular PDUS signal from 0 to 3: • 0 = absence, no synovial flow • 1 = mild, three or fewer isolated signals • 2 = moderate, more than three isolated signals or confluent signal in less than half of the synovial area • 3 = marked, signals in more than half of the synovial area
Karim *et al.* 2004[57]	Semiquantitative assessment for degree of synovitis: • normal = no synovitis • mild = flat, thickened synovium • moderate = thickened synovium with few villi-like protrusions • severe = marked thickening with multiple villi-like protrusions
Schmidt *et al.* 2000[58]	Subjective grading of echogenity of effusion on a 0–3 scale: • 0 = no echoes in the effusion • 1–3 = increasing degrees of echogenity in the effusion PDUS perfusion intensity as analysed subjectively on a 0–3 scale: • 0 = no perfusion • 1–3 = increasing degrees of perfusion
Klauser *et al.* 2004[59]	Criteria for grading intra-articular vascularisation with colour Doppler US/PDUS: • 0 = no intra-articular colour-flow signals • 1 = one to five intra-articular colour-flow signals • 2 = six to 10 intra-articular colour-flow signals • 3 = ≥ 11 intra-articular colour-flow signals

concluded that the standardised definitions produced by OMERACT had improved inter-rater and intra-rater reliability for detecting and grading synovitis of the hand joints.[63] The OMERACT US special interest group defined a Global Synovitis Score, which combines synovial hypertrophy and the PDUS signal in a composite score and has acceptable reliability.[64,65]

Although the 28-joint count is well established for CE, the optimal joint count for US is currently unclear. Joint counts featuring smaller numbers of joints are more feasible for a clinical encounter.[45] Several studies have compared various different joint counts;[45,66–70] however, there is no clear consensus on the optimal joint count at present and joint counts vary between studies, rendering comparison across studies difficult. Research is ongoing.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Patient experience of ultrasound

Ultrasound is relatively inexpensive and safe, avoiding the exposure to radiation that is necessary for conventional radiography and CT. US is a painless and harmless procedure, being non-invasive and not radioactive. As confirmed by this study's patient advisor (see *Appendix 2*), US is comfortable for the patient, unlike MRI, which can be noisy, can make the patient feel claustrophobic and requires lying still for a prolonged period of time. It should be noted that only one patient advisor contributed to this report.

The advent of portable US machines means that scans can be carried out at the bedside or in the outpatient clinic without the need for a second appointment in the radiology department. Most scans for RA involve the hand or wrist joints or focus on the most affected joint and take < 20 minutes.

Current usage in the NHS

Few published data were identified on the current usage of US for RA in the NHS. Therefore, a survey was conducted to address this issue (see *Appendix 1*).

A published survey conducted in 2005[71] analysed data from 126 UK rheumatologists. It found that 93% (117/126) of the surveyed rheumatologists used US results in patient management, with 33% (41/126) conducting the US examination themselves. Of the 60% (76/126) who referred patients to other departments for US, the majority referred patients to the radiology department (56%, 71/126). Synovitis was the second most common indication (estimate from graph: 87%, 110/126) for investigation by US, after tenosynovitis. The authors acknowledge the possibility that US use was overestimated because of the sample of rheumatologists being targeted at imaging events and because of response bias (126 questionnaires returned out of 250 circulated).

Another survey of 100 rheumatologists, conducted in 2009 (Richard Wakefield, University of Leeds, 22 March 2016, personal communication), found that most used US results to detect inflammatory arthritis (84%), but fewer used US for prognostic assessment (44%) or disease monitoring (32%). In total, 28% had conducted the US examination themselves and 81% had referred patients to the radiology department. Of those surveyed, 40% reported that US frequently influenced patient management (Richard Wakefield, personal communication).

The Targeted Ultrasound Initiative,[72] founded in 2012, provides training and resources regarding US. A survey of its international users suggests that > 90% use US for diagnosis and the assessment of remission and approximately 80% use US for routine monitoring (Richard Wakefield, personal communication).

Anticipated costs associated with the intervention

The cost of monitoring synovitis using US was estimated using data provided by our clinical experts (Professor Philip Conaghan, University of Leeds; Dr Cristina Estrach, Aintree University Hospitals NHS Foundation Trust; Professor Christopher Edwards, University of Southampton; and Dr Richard Wakefield, University of Leeds) and a survey (see *Appendix 1*) sent to members of the BSR. Our clinical advisors stated that US used for monitoring synovitis would be undertaken as an outpatient appointment and that contrast would not be required. A broad consensus was reached that two-thirds of scans would take < 20 minutes with one-third taking ≥ 20 minutes. Based on this advice, it was assumed that two-thirds of scans would be costed as NHS reference cost RD40Z (£55.03) and the remainder would be costed as NHS reference cost RD42Z (£59.90), resulting in a weighted cost per scan of £56.66.[73] Therefore, if it is assumed that, on average, four USs are performed on a patient every year, this would equate to a yearly cost of US monitoring of £226.62. No distinction has been made between different types of US for two reasons. First, detailed data on the differential costs have not been identified. Second, there was no clear evidence of a distinction between different US modes regarding the association with synovitis levels detected and in the successful tapering of treatment. Four US scans were chosen as the base case as it was assumed that clinicians may actively monitor patients in whom they had tapered treatment. This assumption was tested in sensitivity analyses.

Chapter 2 Definition of the decision problem

Decision problem

Purpose of the decision to be made

The aim of this assessment was to systematically review the evidence on the use of US examination in addition to clinical assessment, compared with CE only, in relation to how synovitis in patients with RA can be used to predict prognosis or response to treatment and to investigate the cost-effectiveness of US joint examination for monitoring synovitis in terms of guiding decisions about when treatment can be tapered or which patients should be progressed to more intensive treatment.

Intervention

The included intervention was US examination of joints in patients with RA used in addition to CE to detect synovitis. US technologies included were GSUS and PDUS.

Population/setting

The population was adult patients with a confirmed diagnosis of RA at any point in the disease pathway. The setting was secondary care.

Relevant comparators

The comparator was assessment of synovitis by CE, without the use of US technology. This included assessment of inflammatory biomarkers[68] such as ESR or CRP[15] and the use of disease activity scoring tools[68] such as the DAS28,[17] Health Assessment Questionnaire (HAQ),[16] SDAI[18] and Clinical Disease Activity Index (CDAI)[18] (see *Table 1*).

Outside the scope of the decision problem

Ultrasound used in the initial diagnosis of arthritis or to determine the type of arthritis or US to guide injections.

Key factors to be addressed

The review aimed to address the clinical value of US in addition to CE for detecting synovitis at different joints and at different points in the disease pathway, compared with CE alone. The review aimed to investigate the economic costs and benefits associated with the use of US to monitor synovitis in RA.

Overall aims and objectives of the assessment

The review aimed to:

- investigate reported data on the detection of synovitis in RA patients by US and CE and the ability of US and clinically detected synovitis to predict clinical outcome or response to RA treatment
- investigate the economic costs and benefits associated with the use of US to monitor synovitis in RA
- suggest key areas for primary research.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 3 Assessment of ultrasound studies

Methods for reviewing ultrasound studies

The search for evidence on US for monitoring synovitis in RA was undertaken systematically following the general principles recommended in the PRISMA statement.[74] The search strategies used are provided in *Appendix 3*.

Identification of studies

The following electronic databases were searched:

- MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid) (1946 to October 2015)
- EMBASE (via Ovid) (1974 to October 2015)
- Cochrane Database of Systematic Reviews (CDSR) (via Wiley Online Library) (1996–2015)
- Cochrane Central Register of Controlled Trials (CENTRAL) (via Wiley Online Library) (1898–2015)
- Health Technology Assessment (HTA) database (via Wiley Online Library) (1989–2015)
- Database of Abstracts of Reviews of Effects (DARE) (via Wiley Online Library) (1946–2014)
- NHS Economic Evaluation Database (NHS EED) (via Wiley Online Library) (1968–2014)
- Science Citation Index Expanded (via Web of Science) (1900 to October 2015)
- Science Citation Index and Conference Proceedings Index (via Web of Science) (1900 to October 2015)
- ClinicalTrials.gov (US National Institutes of Health) (October 2015)
- European League Against Rheumatism Abstract Archive (via Web of Science) (October 2015)
- American College of Rheumatology and Association of Rheumatology Health Professionals (via Web of Science) (October 2015)
- OMERACT conference proceedings (via Web of Science) (October 2015).

In addition to electronic database searching, bibliographies of systematic reviews and included studies were checked for studies meeting the review inclusion criteria.

Initial scoping searches identified eight relevant guidelines.[9,24,28,30,46,75–77] Bibliographies of these guidelines were also checked.

Following the submission of the report for peer review, the full results of one of the included studies [Targeting Synovitis in Early Rheumatoid Arthritis (TaSER)] were published.[78] On the request of a peer reviewer, the results from this publication have been included. The searches were not updated. However, as this was a publication of an already included study, with data relevant to the treatment strategies, it was decided that the publication would be included. Additionally, the Aiming for Remission in Rheumatoid Arthritis (ARCTIC) trial, although not published in full at the time of peer review, was published in a preliminary conference abstract in November 2015.[79] It was decided that this publication would also be discussed.

Inclusion and exclusion criteria

Population

Adult patients diagnosed with RA.[11]

Intervention

The intervention was US of joints in patients with RA, as used to assess synovitis. Studies were not excluded on the basis of type of US technology used. US technologies included were GSUS and PDUS.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Comparators

The comparator was assessment of synovitis by CE without the use of US technology (in the same patients). This included assessment of inflammatory biomarkers such as ESR or CRP[15] and the use of disease activity scoring tools such as the DAS28[17] and HAQ.[16]

Outcomes

Outcomes were diagnostic accuracy measured by sensitivity and specificity (CE with reference US, or reported US with reference CE, or sufficient data reported to calculate sensitivity or specificity); responsiveness to change in inflammation as measured by standardised response mean (SRM);[80] synovitis detection rate; prognostic sensitivity or prognosis associated with baseline measures; prediction of response to treatment; or the use of US in clinical decision-making.

For diagnostic accuracy data, studies were accepted if they reported sensitivity or specificity or if they provided sufficient data to calculate sensitivity or specificity.

Study design

Diagnostic studies, prognostic studies and studies investigating prediction of response to treatment, or the influence of US on treatment decisions, whether as a one-off US test or as serial US testing, were included. Studies were not excluded based on quality alone; however, for prognostic studies, the method and setting of outcome measurement had to be the same for all study participants.[81]

Additionally, systematic reviews were sought and were used to identify studies meeting the inclusion criteria for the review.

Exclusion criteria

The exclusion criteria were as follows:

- US used for RA diagnosis only (for initial diagnosis of arthritis or for differentiating between types of arthritis)
- US used to guide injections
- US used for inflammatory conditions other than RA
- studies with healthy control subjects unless outcome data were reported separately for the RA subgroup
- studies of US compared with another imaging technology (or other diagnostic method) for which clinical comparator data were not available or for which data did not allow comparison of US with CE
- studies with a low proportion (< 80%) of patients diagnosed with RA, unless outcome data were reported separately for the RA subgroup
- studies of US reporting inter-rater or intra-rater reliability only
- studies of tenosynovitis not reporting data about synovitis
- animal models
- preclinical and biological studies
- narrative reviews, editorials and opinions
- non-English-language papers
- abstracts reporting insufficient details to allow inclusion.

This review concentrated on US as the intervention. There is no conclusive gold standard/reference standard for assessing synovitis related to disease progression. This review investigated the association of US-detected synovitis with later outcomes in prognostic studies (see *Assessment of prognostic studies*). This can be considered as validating the test results.[28] For detection of synovitis, it may be considered that surgery is the gold standard/reference standard, allowing direct visualisation of the synovial membrane.[57] In clinical practice, this would not be used for monitoring synovitis, and limiting studies to those including surgery would limit the amount of studies eligible for the review. In terms of other imaging techniques that have a potential role in monitoring, it has been suggested that MRI may be used as a gold standard/reference standard.

However, it has been reported that US and MRI detect similar rates of inflammation.[28] Scintigraphy and positron emission tomography (PET) have been reported to detect similar rates of inflammation to CE.[28] Other imaging techniques with a potential for use in RA patients are conventional radiography and CT scans. US is more likely to be practical than other imaging techniques in assessing synovitis. US is more comfortable for the patient than MRI, is less expensive and has an advantage over other imaging techniques as it can be used immediately after CE to assess symptomatic areas.[30] Limiting studies to those including an additional imaging technique would have limited the number of studies eligible for the review. As this review concentrated on US as the intervention, the decision was taken not to compare US-detected synovitis with synovitis detected by other imaging techniques.

Study selection

Titles and abstracts were examined by one reviewer. Study selection based on full texts was carried out by two reviewers, with discrepancies resolved by discussion.

Data extraction strategy

Data relevant to the decision problem were extracted from all studies by one reviewer using a standardised data extraction form. All data extracted were checked thoroughly by a second reviewer. Data were extracted without blinding to authors or journal. Discrepancies were resolved by discussion.

Quality assessment strategy

The methodological quality of each included study was assessed by one reviewer and checked by a second reviewer. Discrepancies were resolved by discussion. Diagnostic studies were assessed using criteria based on the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) tool.[82] The QUADAS tool was selected as it is a validated, evidence-based tool that has been widely used in diagnostic review. Prognostic studies were assessed using criteria taken from the Generic Appraisal Tool for Epidemiology (GATE)[83] and the Quality in Prognosis Studies (QUIPS)[81] tool. These are both validated tools that have been used for reviewing prognostic studies. Two items of particular relevance to RA studies were adapted from recommendations on inclusion criteria and study design for RA studies by Karsh *et al.*[84]

Methods of analysis/synthesis

For calculations of diagnostic accuracy, US was counted as the reference standard and the accuracy of the clinical comparator was assessed using sensitivity, that is, the proportion of true positives (TPs), and specificity, that is, the proportion of true negatives (TNs), calculated as follows: sensitivity is the number of TPs divided by the sum of the TPs and the false negatives (FNs); specificity is the number of TNs divided by the sum of the TNs and the false positives (FPs).

Data were tabulated and discussed. Evidence synthesis would be attempted unless precluded by heterogeneity (of population, intervention, comparator or outcomes).

Survey

In addition to the systematic review, a survey was conducted to investigate current usage of US in treatment decisions for RA patients. The survey was publicised by the BSR through newsletters. The questions used in the survey are provided in *Appendix 1*. The survey was available to clinicians from December 2015 to February 2016.

Results

Quantity and quality of research available

A total of 2724 records were identified from the electronic databases following deletion of duplicate records.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Eight guidelines were identified through grey literature searching.[9,24,28,30,46,75–77] Eight systematic reviews of relevance to the decision problem were also identified.[25,29,45,50,85–88] Guidelines were published on US in rheumatology by the ACR,[46] Arthritis Research UK[30] and EULAR,[75] which also published guidelines on imaging in RA.[28] Guidelines on the management of RA were published by NICE[9,77] and the Scottish Intercollegiate Guidelines Network.[76] Reviews were published on the use of US in inflammatory arthritis[29,85–87] or US for use in RA including the detection of synovitis and defining remission.[24,25,45,50,88] The bibliographies of these guidelines and reviews were searched to identify studies meeting the review inclusion criteria, identifying an additional 26 records. This led to a total of 2750 records being screened.

The study selection process is provided in *Figure 1*. At title and abstract sift, 2596 records were excluded. The 63 records excluded at full paper sift are presented with reasons for exclusion in *Appendix 4*.

A total of 75 articles[53,55,68,69,89–158] describing 58 studies were included in the review; following the searches but prior to publication of this report, one of the studies identified as ongoing by the searches (ARCTIC) was published as an abstract.[79] Data extraction forms for these included studies are provided in *Appendix 5*.

Study quality assessment forms for the included studies are provided in *Appendix 6*.

Ten[89–98] of the 58 included studies were reported as abstracts only and so limited details were available.

Of the 58 studies included, diagnosis of RA was reported as being confirmed using the ACR or EULAR classification criteria for RA, either the 2010 criteria[11] or an earlier version,[8] in all except nine studies.[89,90,92–96,98,99]

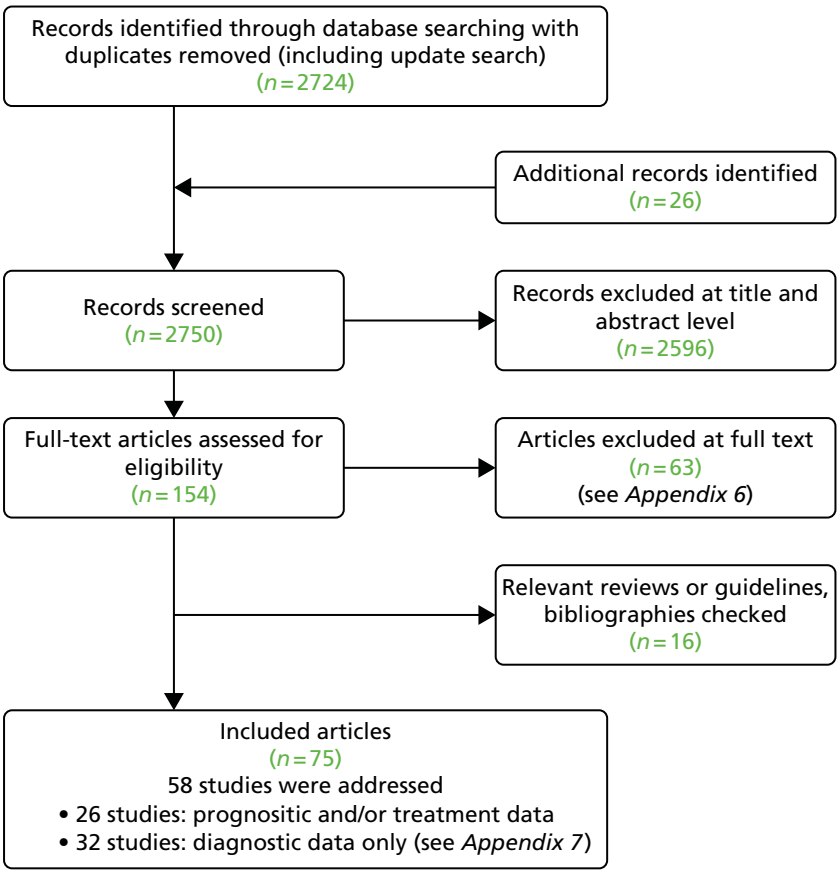


FIGURE 1 Flow diagram of study inclusion.

An established US scoring system was referenced in all but nine[89,91,94,95,100–104] of the 58 included studies. According to the hierarchy of evidence published by Merlin *et al.*,[159] diagnostic accuracy studies are of higher quality if they are blinded and recruit consecutive patients and prospective cohort studies are of higher quality than other forms of prognostic studies. Of 33 studies[53,92,96,102–116,118–132] reporting diagnostic accuracy or detection rates of synovitis, 13[102–104,107,109–111,113–115,118,119,124] were blinded comparisons of consecutive patients. Of 15 studies of prognosis,[55,69,97,98,113,134,135,137–144] all but one,[138] an ancillary study to a RCT, were prospective cohort studies.

Studies reporting the detection of synovitis are described in *Appendix 7* (see *Tables 138–140*). Diagnostic data were extracted from 37 publications[53,68,92,96,99,102–132,146] reporting on 33 studies. One of these studies[113,146] also provided prognostic data.

For most of the studies providing diagnostic data (see *Appendix 7*, *Table 138*), CE had a high specificity and low sensitivity when using PDUS or GSUS as the reference standard. This indicates some agreement between CE and US, with US detecting synovitis in some joints that CE did not detect and only a few cases in which CE detected synovitis and US did not. This agrees with the higher detected rates of synovitis for US over CE (see *Appendix 7*, *Table 139*) reported in the majority of studies.

However, five studies found lower rates of detection of synovitis with US than with CE.[106,115,123,125,130] Also, there were mixed results in three studies, with higher detection rates for GSUS than for CE but lower detection rates for PDUS than for CE in two studies[108,119] and lower detection rates for PDUS than for CE for metacarpophalangeal (MCP) joints but higher detection rates for PDUS than for CE for proximal interphalangeal (PIP) and wrist joints in one study.[131]

The five studies[106,115,123,125,130] that found lower rates of detection with US than with CE were some of the older trials; however, there were other trials from the early 2000s that did not report this finding and so this cannot explain the difference between these trials and trials finding a higher rate of detection with US than with CE. Study quality did not explain the differences between the study results, although one of the studies reporting a lower rate of detection with US than with CE had a small sample size ($n = 6$), and the differences did not appear to be caused by the US scoring system used (0–1 vs. 0–3 or 0–4; see *Table 3*).

It may be that the types of joint examined could explain the differences between the study results. Of the studies finding lower rates of detection with US than with CE, three studies[106,115,130] investigated ankle and/or foot joints and one study[123] included metatarsophalangeal (MTP) joints; however, this study also assessed PIP and MCP joints, which have been shown to have higher rates of detection of synovitis using US than using CE.[125] This agrees with one within-study comparison of joints, which reported that CE had a lower sensitivity for MTP joints than for MCP or PIP joints.[160] However, heterogeneity between other studies, suggesting that US is less useful for foot and ankle joints than for wrist and hand joints, means that there is uncertainty in the findings.

The detection of subclinical synovitis would be useful only if clinically relevant, as investigated using prognostic studies.

Prognostic data were extracted from 23 publications[55,56,69,97,98,100,113,134–139,141–148,150,151] reporting on 15 studies. One of these studies was an ancillary study to a RCT.[138] The other studies were prospective cohort studies in which a cohort of patients was assessed at baseline with US and with CE and was then followed up to assess the association of these baseline measures with a given outcome. Prognostic data are provided in *Tables 6–8*.

Treatment data were extracted from 16 publications[89–91,93–95,100,101,139,152–157] reporting on 13 studies. Following the searches but prior to publication of this report, one of the studies identified as ongoing (ARCTIC) was published as an abstract.[79] This abstract was included to provide treatment data, meaning that 14 studies in total provided data relating to treatment. Six of these were prospective cohort studies that measured the association of an outcome at follow-up with the US and CE variables measured at baseline[93,101] or US and CE variables measured

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

after 4 months of treatment[139] or at the time of treatment discontinuation or tapering.[95,155,156] One study reported data from a prospective cohort in which baseline US and CE variables were tested for association with an outcome taken from a RCT of RA treatment.[95,132] Two studies were RCTs comparing treatment strategies with or without US.[79,154] One of these RCTs[154] also provided data on treatment decisions. Treatment decisions were also investigated using observational data.[89–91,94,152] Treatment data are provided in *Tables 9–14*.

Meta-analyses were not performed because of heterogeneity in the type of US used, US scoring systems used (see *Table 3*), joints assessed, clinical comparators (type of examination; use of composite clinical scoring), outcome measures and follow-up times.

Assessment of prognostic studies

Heterogeneity of trials precluded meta-analysis. Therefore, no summary estimates of effect were available, which is a limitation of the review. Significance values refer to the association of baseline US and CE measures with later outcome measures. Radiographic progression measures are of importance to patients as they reflect structural damage, which is associated with loss of function over time. Other outcomes assessed are of importance to patients as they reflect disease activity and function. Studies investigated whether or not baseline measures could be predictive of later outcomes, rather than directly comparing different baseline measures. As well as heterogeneity of outcome measures, studies differed in the statistics reported to assess the association of baseline US and CE measures with clinical outcome at follow-up, including Spearman correlations (*R*) and odds ratios (ORs) (univariate analysis unless otherwise stated).

Two studies[113,134] reported the sensitivity of US and CE measures to predict progression (*Table 4*). These were prospective cohort studies in which patients were tested with US and underwent CE at baseline and were then followed up and evaluated for the outcome measures of progression. Boyesen and Haavardsholm[134] used a measure of GSUS inflammation based on synovitis and tenosynovitis and found that this US measure had a similar sensitivity for predicting erosive progression as MRI at 12 months, indicating that the majority of patients with a high US score, or high DAS, relapsed. This measure of GSUS inflammation[134] had a higher specificity than DAS28 (although neither US nor DAS28 had high specificity) for predicting erosive progression as measured by MRI at 12 months, indicating that more patients with a low GSUS score than a low DAS did not progress. Ikeda *et al.*[113] found that GSUS, PDUS and Disease Activity Score 28 joints using C-reactive protein (DAS28-CRP; see *Table 1*) had similar sensitivities for predicting radiographic progression at 24 weeks (see *Table 4*). GSUS had a lower specificity than PDUS and DAS28 for predicting radiographic progression.[113] This study subgrouped patients by recently retrieved treatment and found that in MTX-treated patients ($n = 16$) and TCZ-treated patients ($n = 17$) PDUS had higher specificity than DAS28-CRP, but that specificities were similar in TNFi-treated patients ($n = 24$) (see *Appendix 5*). Both of these studies were prospective cohort studies, at level II in the study quality hierarchy, although the study by Ikeda *et al.*[113] was a better-quality and more rigorous study given that blinding was unclear in the study by Boyesen and Haavardsholm[134] (see *Appendix 8*).

Eleven studies (with 14 publications[55,69,97,113,134–136,138–141,144,148,150]) reported an association of US and CE prognostic factors with radiographic progression/erosion (*Table 5*). These were prospective cohort studies in which patients were tested with US and underwent CE at baseline and were then followed up and evaluated for the outcome measures of progression.

Of the 11 studies, seven[55,69,113,138–140,144] found that US and at least one of the clinical measures investigated could significantly predict radiographic progression, most of which were of high quality (with the exception of two studies[69,138] in which blinding was unclear). Three studies found that US could significantly predict radiographic progression whereas the clinical comparator could not (DAS28-ESR/ESR/CRP;[134] joint counts;[135] SJC[97]), two[97,134] of which were less rigorous in terms of study quality, with blinding being unclear. One study[141] found that neither US nor the clinical comparator (ESR) was significantly associated with later progression; this study was more rigorous in terms of study quality. All of the studies were prospective cohort studies, at level II in the study quality hierarchy, with the exception of one study,[138] which was an ancillary study to a RCT (see *Appendix 8*).

TABLE 4 Sensitivity and clinical prognosis

Study	Population	Outcome	Patients with outcome, *n* (%)	US measure	Sensitivity (95% CI), %	Specificity (95% CI), %	Clinical measure	Sensitivity, %	Specificity, %
Boyesen 2011[134]	79 patients, dominant wrist	MRI erosive progression at 12 months (≥ 1-unit increase in RAMRIS)	53 (67.1)	GSUS inflammation score cut-off values < 0.5 vs. ≥ 0.5[a]	79 (NR)	55 (NR)	DAS28 remission/low vs. moderate/high[b]	74 (NR)	21 (NR)
Ikeda 2013[113]	57 patients, 28 joints of DAS28	Radiographic progression (mTSS) at 24 weeks[c]	21 (36.8)	Total GSUS score cut-off values of < 62	56 (NR)	57 (NR)	DAS28-CRP cut-off of < 9.0	64 (NR)	81 (NR)
				Total PDUS score cut-off values of < 21	69 (NR)	76 (NR)			

CI, confidence interval; mTSS, modified total Sharp score; NR, not reported; RAMRIS, rheumatoid arthritis magnetic resonance imaging scoring system.
a Composite score of synovitis and tenosynovitis.
b Cut-off values not reported.
c vdHSS.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 5 Correlation of US and CE with prognosis: radiographic progression

Study	Population	Joints	Follow-up	Outcome	Type of US	Correlation of US synovitis with outcome	Type of CE	Correlation of CE with outcome
Backhaus 2013[69]	453 patients	US7 score	12 months	Bone erosion (US7 erosion sum score)	US7 synovitis sum score in GSUS	$R \pm$ SD: 0.140 ± 0.038 (95% CI 0.066 to 0.215; $p < 0.001$)	DAS28, ESR, CRP	$R \pm$ SD: DAS28 of 0.621 ± 0.141 (95% CI 0.345 to 0.898; $p < 0.001$); ESR 0.008 ± 0.010 (95% CI −0.011 to 0.027; $p = 0.399$); CRP 0.001 ± 0.009 (95% CI −0.017 to 0.018; $p = 0.945$)
					US7 synovitis score in PDUS	$R \pm$ SD: 0.099 ± 0.051 (95% CI −0.001 to 0.200; $p = 0.053$)		
Boyesen 2011[134]	84 patients	Dominant wrist	1 year	MRI erosive progression[a]	GSUS (univariate analysis)[b]	OR 2.15 (95% CI 1.23 to 3.75; $p = 0.01$)	DAS28 (univariate analysis)	OR 1.15 [95% CI 0.80 to 1.64; $p = 0.46$ (NS)]
					GSUS (multivariate analysis)[b,c]	OR 2.01 (95% CI 1.14 to 3.53; $p = 0.02$)	ESR (univariate analysis)	OR 1.02 [95% CI 0.98 to 1.06; $p = 0.29$ (NS)]
					GSUS synovitis (radiocarpal joint assessed in the dorsal midline)	OR 7.2 (95% CI 0.9 to 61.0; $p =$ NS)	CRP (univariate analysis)	OR 1.04 [95% CI 0.98 to 1.10; $p = 0.16$ (NS)]
Brown 2008[135] (same study as Ikeda 2007[136])	90 patients in clinical remission	Dominant hand MCP joint and wrist	12 months	Progressive radiographic damage (both hands and feet)[d]	PDUS signal	OR 12.21 (95% CI 3.34 to 44.73; $p < 0.001$)	Painful joint count	OR 3.32 (95% CI 0.39 to 28.30; $p = 0.273$)
					GSUS synovial hypertrophy	OR 1.92 (95% CI 0.49 to 7.54; $p = 0.350$)	TJC	OR 2.17 (95% CI 0.26 to 18.10; $p = 0.473$)
							SJC	OR 1.69 (95% CI 0.43 to 6.67; $p = 0.457$)
Ikeda 2007[136] (same study as Brown 2008[135])	107 patients in remission (clinician assessed) on DMARDs	Dominant hand MCP and wrist	12 months	Radiographic progression (direct assessment of individual joints)	GSUS synovial hypertrophy score	Mann–Whitney U-test $p = 0.027$	Painful joints	Chi-squared test $p = 0.284$
					PDUS score	$p < 0.001$ (in asymptomatic joints only $p = 0.002$)	Tender joints	$p = 0.389$
					Increased PDUS signal	Likelihood ratio 7.02, chi-squared test $p = 0.037$	Swollen joints	$p = 0.417$

Study	Population	Joints	Follow-up	Outcome	Type of US	Correlation of US synovitis with outcome	Type of CE	Correlation of CE with outcome
Cavet 2009[138]	24 patients	10 MCP joints[100]	110 weeks	Radiographic progression[e]	GSUS synovial thickening	$R = 0.59$ (p-value NR)	Biomarkers (93 serum proteins) at 6 weeks (multivariate model)	Significant correlation $R = 0.79$ (p-value NR)
					Power Doppler area)	$R = 0.77$ (p-value NR)		
Dougados 2013[139,149]	59 patients	MCP (×10), PIP (×10), wrist (×2) and MTP (×10) joints	2 years	Radiological progression[f]	GSUS (0 vs. 1–3)	OR 1.64 (95% CI 1.08 to 2.47)	Clinical synovitis[g]	0 vs. 1–3: OR 2.08 (95% CI 1.39 to 3.11; $p < 0.001$); 0–1 vs. 2–3: OR 1.64 (95% CI 1.14 to 2.36; $p = 0.008$)
					GSUS (0–1 vs. 2–3)	OR 1.91 (95% CI 1.18 to 3.10)	Tender joints	OR 1.53 (95% CI 1.02 to 2.29; $p = 0.04$)
					PDUS 0 vs. 1–3	OR 1.80 (95% CI 1.20 to 2.71; $p = 0.005$)	Tender joints with clinical synovitis	OR 1.89 (95% CI 1.25 to 2.85; $p = 0.002$)
					PDUS 0–1 vs. 2–3	OR 1.36 (95% CI 0.78 to 2.38; $p = 0.278$)		
Ikeda 2013[113]	57 patients (MTX, $n = 16$; TNFi, $n = 24$; TCZ, $n = 17$)	Finger, wrist, elbow, shoulder and knee	24 weeks	Radiographic damage[e]	GSUS (cumulative total GSUS score)	$R = 0.062$; $p = 0.649$ (MTX, $r = 0.210$, $p = 0.436$; TNFi, $r = 0.027$, $p = 0.900$; TCZ, $r = 0.492$, $p = 0.179$)	Cumulative DAS28-CRP	$R = 0.342$, $p = 0.009$ (MTX, $r = 0.487$, $p = 0.056$; TNFi, $r = 0.308$, $p = 0.144$; TCZ, $r = 0.369$, $p = 0.145$)
					PDUS (cumulative total PDUS score)	$R = 0.357$; $p = 0.006$ (MTX, $r = 0.679$, $p = 0.004$; TNFi, $r = 0.153$, $p = 0.476$; TCZ, $r = 0.353$, $p = 0.165$)		
Naredo 2007[55]	42 RA patients starting DMARDs (38 patients followed up to 1 year)	28 joints	1 year	Total radiographic score progression (radiographic erosion score and JSN score)[e]	Joint count for GSUS[h]	GSUS $r = 0.61$; $p < 0.001$	SJC, TJC, DAS28, HAQ, ESR, CRP	SJC, $r = 0.46$, $p < 0.01$; TJC, $r = 0.36$, $p < 0.05$; DAS28, $r = 0.40$, $p < 0.05$; HAQ, $r = 0.36$, $p < 0.05$; ESR/CRP, NS
					Joint count for PDUS signal[i]	PDUS $r = 0.59$; $p < 0.001$		

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 5 Correlation of US and CE with prognosis: radiographic progression (*continued*)

Study	Population	Joints	Follow-up	Outcome	Type of US	Correlation of US synovitis with outcome	Type of CE	Correlation of CE with outcome
Naredo 2008[140]	278 patients with complete data (of 367 RA patients starting TNFis)	86 intra-articular and periarticular sites in 28 joints	1 year	Total radiographic score progression (radiographic erosion score and JSN score);[e] multivariate analysis[j]	Time integrated values of US joint count for PDUS signal	Total radiographic score ($R = 0.59$) significant; erosion score ($R = 0.64$) significant; JSN score NS	ESR, CRP	Total radiographic score: ESR ($R = 0.59$) significant; ESR for erosion or JSN score NS; baseline CRP level weak predictive value for JSN score progression ($R = 0.2$); CRP NS for erosion and total scores
Osipyants 2013[97,150]	35 patients	Wrists	1 year	Radiographic progression (mTSS)[e]	PDUS signals of residual inflammation of the wrists at 6 months	Significantly correlated ($r = 0.696$; $p = 0.011$); mTSS at 1 year significantly higher [93 (range 56–131) vs. 32.5 (range 19.5–83) units; $p < 0.04$] in patients with middle- or high-active PDUS signals ($n = 27$) (SJC > 1, PDUS > 1) than in cases with lower imaging activity or the absence of the PD signals ($n = 8$) at 6 months	SJC	Non-significantly lower radiographic progression over 1 year in those with SJC ≤ 1 [55 (range 31–116) vs. 88.5 (range 55–130); $p < 0.205$] than in those with SJC > 1 ($n = 25$)
Reynolds 2009[141]	25 patients (providing data at the 2-year follow-up of 40 patients recruited)	MCP and PIP joints	2 years	Erosion progression[k]	GSUS and PDUS	Mann–Whitney *U*-test synovial score (0–3) NS; mean synovial thickness (cm) NS; pre-contrast PDUS score (0–3) NS; post-contrast PDUS score (0–3) NS	ESR	NS

Study	Population	Joints	Follow-up	Outcome	Type of US	Correlation of US synovitis with outcome	Type of CE	Correlation of CE with outcome
Yoshimi 2013[144]	22 RA patients in clinical remission (of 31 recruited)	Bilateral wrists and all of the MCP and PIP joints	2 years	Radiographic progression[e]	Total PDUS score	Total PDUS score significantly higher in patients with radiographic progression than in patients without (6.00 ± 6.44 vs. 0.87 ± 1.15; $p = 0.0011$)	SJC, DAS28, gVAS and the serum levels of ESR, CRP and MMP-3	Significant differences in SJC (0.33 ± 0.79 vs. 1.29 ± 0.70; $p = 0.0040$) and DAS28-ESR ($p = 0.0010$) and DAS28-CRP ($p = 0.041$) between the radiographic progression group and the non-progression group. Borderline significance for TJC ($p = 0.054$). No significant difference in other clinical parameters (gVAS and the serum levels of ESR, CRP and MMP-3) between the radiographic progression group and the non-progression group
					Total GSUS score	No significant difference in the total GSUS score, (12.6 ± 12.4 vs. 8.80 ± 5.78; $p = 0.57$), between the radiographic progression group and the non-progression group		

CI, confidence interval; gVAS, general visual analogue scale; JSN, joint space narrowing; MMP-3, matrix metalloproteinase 3; mTSS, modified total Sharp score; NR, not reported; NS, non-significant radiographic progression; PD, power doppler; RAMRIS, rheumatoid arthritis magnetic resonance imaging scoring system; SD, standard deviation; US7, ultrasound 7 score.

a ≥ 1-unit increase in 1-year RAMRIS erosive change.
b Composite score of synovitis and tenosynovitis.
c CE measures were not included in the stepwise regression analysis as they did not reach statistical significance. Multivariate analysis included GSUS, bone marrow oedema, gender and age.
d Genant-mTSS.
e vdHSS.
f Occurrence or worsening of erosion or JSN.
g 0 = no synovitis; 1 = doubtful synovitis; 2 = obvious and moderate synovitis; 3 = obvious and important synovitis.
h Grey-scale USJCAS (ultrasonographic joint count for active synovitis).
i Power Doppler USJIPD (ultrasonographic joint index for power Doppler signal).
j Multivariate analysis included time-integrated PDUS, rheumatoid factor, CRP and ESR.
k Measured by US, graded from 0 to 3.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The majority of the studies reported a significant correlation between US and radiographic progression, using GSUS[55,69,134–136,138,139,148] or PDUS;[55,69,97,113,134–136,138–140,144,148] most of these studies were rigorous in terms of study quality, with four less rigorous studies in which blinding was unclear.[69,97,134,138]

In some cases, GSUS was not significantly correlated with radiographic progression whereas PDUS was (see *Table 5*). There was no significant correlation with radiological progression for GSUS in the study by Ikeda *et al.*,[113] which did not find improved sensitivity or specificity above those for DAS28-CRP at 24 weeks (see *Table 4*). This study also found that PDUS was significantly correlated with radiological progression in the MTX-treated subgroup but not in bDMARD-treated patients.

Grey-scale ultrasound was not significantly correlated with radiographic progression in a study of 22 patients in clinical remission whereas there was a significant correlation between PDUS and radiographic progression.[144] One study[135,136] found that PDUS was significantly associated with later radiographic outcomes. Another study of 25 patients found that neither GSUS nor PDUS was significantly correlated with erosion progression at 2 years' follow-up.[141]

Some studies reported a significant association of CE with radiographic progression (see *Table 5*). DAS28 was significantly correlated with radiographic progression in four studies,[55,69,113,144] although not in the study by Boyesen and Haavardsholm[134] nor in the bDMARD subsets in the study by Ikeda *et al.*[113] HAQ score was significantly correlated with radiographic progression in the only study reporting this outcome.[55] ESR and CRP did not significantly correlate with radiographic progression,[55,69,134,141,144] with the exception of the study by Naredo *et al.*,[140] which found that ESR and total radiographic score progression were significantly associated. Biomarkers (serum proteins) were significantly correlated with modified total Sharp score (mTSS) in the only study reporting this comparator.[138]

Swollen joint count and TJC were significantly correlated with radiographic progression in three studies[55,139,144] but not in two studies[97,135,136] and painful joint count was not significantly correlated with radiographic progression in one study.[135,136]

Power Doppler ultrasound was significantly associated with radiographic progression in 10 studies in which it was measured. Associations reported as ORs were as follows: OR 12.21 [95% confidence interval (CI) 3.34 to 44.73; $p < 0.001$][135] and OR 1.80 (95% CI 1.20 to 2.71; $p = 0.005$).[139] Associations reported at correlation coefficients were as follows: $r = 0.099$ ($p = 0.05$);[69] $r = 0.357$ ($p = 0.006$);[113] $r = 0.59$ ($p < 0.001$);[55] $r = 0.696$ ($p = 0.011$);[97] $r = 0.59$ (p-value not reported);[140] and $r = 0.77$ (p-value not reported).[138] Mann–Whitney U-test results reported were $p < 0.001$[136] and $p = 0.0011$.[144]

One study[141] reported a non-significant association of baseline PDUS with the outcome measure. This different result could not be explained by study quality or the joints assessed and, although the authors suggested that the scoring system used was not as sensitive as that used in other studies, other studies that used semiquantitative scores found significant associations. This study did differ from most in that the outcome of erosion was measured by US. The other study that used US to measure the end point was the study reporting borderline significance for association.[69]

Grey-scale ultrasound was significantly associated with radiographic progression in six studies. Significant associations reported were as follows: Mann–Whitney U-test $p = 0.027$,[136] $r = 0.140$ ($p < 0.001$),[69] $r = 0.61$ ($p < 0.001$),[55] OR 2.15 (95% CI 1.23 to 3.75; $p = 0.01$)[134] and OR 2.08 (95% CI 1.39 to 3.11).[139] A moderate correlation was found between GSUS and radiographic progression in one study.[138] GSUS was not significantly associated with radiographic progression in three studies: OR 1.92 (95% CI 0.49 to 7.54; $p = 0.350$);[135] $R = 0.062$ ($p = 0.649$);[113] and Mann–Whitney U-test non-significant.[141] One of these studies[113] had a shorter follow-up time (24 weeks) than the others, but the other two studies[135,141] with non-significant associations had similar follow-up times to the studies finding significant associations. These two studies[135,141] were two of the older studies included in the review, which may indicate a difference in the US machines used; however, because of the heterogeneity between studies, it is uncertain why these studies differed from those reporting

a significant association. The difference between studies reporting significant associations and studies reporting non-significant associations could not be explained by study quality, joints assessed or how the end point was measured.

The DAS28 joints was significantly correlated with radiographic progression in one study measuring progression using the ultrasound 7 score (US7) score ($r = 0.621$, $p < 0.001$)[69] and three studies measuring progression using the mTSS [$r = 0.342$, $p = 0.009$;[113] $r = 0.40$, $p < 0.05$;[55] and $p = 0.0010$ (DAS28-ESR) and $p = 0.041$ (DAS28-CRP) (only p-values reported)[144]]. DAS28 was not significantly correlated in one study measuring MRI erosive progression.[134] Given heterogeneity between trials, it is uncertain whether outcome measure was the factor explaining the difference between trials finding a significant association, or not, between DAS28 and later radiographic progression.

Other outcomes

Studies reporting outcomes other than radiographic progression (measures of disease activity) are shown in *Table 6*. These were prospective cohort studies in which patients were tested with US and underwent CE at baseline and were then followed up and evaluated for outcome measures of disease activity. All studies used measures in the baseline CE that were not the same as the outcome measure at follow-up (although some studies additionally reported assessment of the outcome measure at baseline (DAS28 and HAQ score;[55] components of the composite outcome measure[145]).

In three cases, US and at least one clinical comparator measure were associated with the later outcome [ESR,[98] Health Assessment Questionnaire Disability Index (HAQ-DI)[142] and DAS[55]]. Three studies found that US was significantly associated with the later outcome but the clinical comparator was not (DAS28-CRP or SDAI;[97] DAS28-CRP, DAS28-ESR, SJC, TJC, CRP and ESR;[145] and SJC and DAS28[137]). In the study by Naredo *et al.*,[55] the clinical comparator DAS28 was associated with the later HAQ outcome, but the US joint count of 28 joints was not. The authors suggested that their results may differ from those of other studies because early RA patients were used in their study and, at this point of disease, functional status is related to inflammatory activity more than residual structural damage.

Low disease activity (i.e. a DAS of < 2.4) at 1 year was not significantly correlated with PDUS or the biomarker chemokine (C-X-C motif) ligand 13 (CXCL13).[147]

The Disease Activity Score 28 joints at 1 year was significantly correlated with GSUS and PDUS and HAQ, SJC, TJC and ESR/CRP.[55] Relapse from clinical remission (i.e. of DAS28 of < 2.6) was significantly correlated with PDUS and ESR but not transforming growth factor beta 1 (TGFβ1).[98] Relapse (i.e. a DAS of ≥ 1.6) at 6 months was significantly correlated with PDUS, SJC and DAS but not GSUS.[137] Time to clinical remission (DAS28 of < 2.6) was not significantly correlated with GSUS, PDUS or DAS28 at baseline in a study of 10 patients.[143]

Remission, defined by ACR/EULAR Boolean criteria [score of ≤ 1 for all of the TJC, SJC, CRP and patient global assessment components], at 2 years was significantly correlated with PDUS and GSUS, but not with the clinical comparators DAS28, SJC, TJC, CRP or ESR.[145]

Flare, defined as any increase in disease activity requiring a change in therapy, at 12 months was significantly correlated with PDUS, HAQ and Rheumatoid Arthritis Quality of Life (RAQoL), but not with GSUS, SJC, TJC, CRP, DAS28 or SDAI.[142]

Two studies investigated HAQ, one of which found that HAQ score at 1 year of follow-up was not significantly correlated with GSUS, PDUS, SJC, or ESR, but was with DAS28 and TJC.[55] One study[97] found that HAQ score at 6 months was significantly correlated with PDUS, but not with DAS28 or SDAI (DAS28 and SDAI were significantly correlated in the subgroup with a disease duration of < 2 years but there were only nine patients in this subgroup).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 6 Correlation of US and CE with prognosis outcomes other than radiographic progression

Study	Population	Joints	Follow-up	Outcome	Type of US	Correlation of US synovitis with outcome	Type of CE	Correlation of CE with outcome
Scirè 2009[137] (same study as Bugatti 2012[147])	43 patients achieving clinical remission	Bilateral shoulder, elbow, wrist, MCP, PIP, sternoclavicular, acromioclavicular, knee, ankle and MTP joints	6 months	Clinical relapse	US joint count > 2 (multivariable)[a]	OR 4.6 (95% CI 0.4 to 49.5; NS)	SJC > 1	OR 0.6 (95% CI 0.1 to 5.5)
					PDUS > 0	OR 12.8 (95% CI 1.6, 103.5; $p < 0.05$)	DAS28 of > 1.1	OR 9 (95% CI 0.7 to 110.3)
Bugatti 2012[147] (same study as Scirè 2009[137])	161 patients with early RA	Hands	12 months	Low disease activity (i.e. a DAS of < 2.4)	PDUS synovitis scored (0–3)	OR 0.94 [0.31 to 2.83; $p = 0.91$ (NS)]	Serum levels of the chemokine CXCL13	NS
Naredo 2007[55]	42 patients with RA starting DMARDs (38 followed up for 1 year)	28 joints	1 year	DAS28	Joint count for GSUS[b]	GSUS $r = 0.63$; $p < 0.001$	SJC, TJC, DAS28, HAQ, ESR, CRP	DAS28, $r = 0.75$, $p < 0.001$; HAQ; $r = 0.66$, $p < 0.001$; TJC, $r = 0.50$, $p < 0.01$; SJC; $r = 0.45$, $p < 0.01$; ESR/CRP, $r = 0.49$, $p < 0.01$
					Joint count for PDUS signal[c]	PDUS $r = 0.63$; $p < 0.001$		
				HAQ	Joint count for GSUS[b]	GSUS NS	SJC, TJC, DAS28, HAQ, ESR, CRP	HAQ, $r = 0.82$, $p < 0.001$; DAS28; $r = 0.49$, $p < 0.01$; TJC, $r = 0.50$, $p < 0.01$; SJC NS; ESR NS
					Joint count for PDUS signal[c]	PDUS NS		
Osipyants 2013[97,150]	35	Wrists	6 months	HAQ	PDUS	Significant ($r = -0.821$; $p = 0.003$)	DAS28-CRP, SDAI	Association with baseline DAS28 or SDAI NS. Subgroup [disease duration < 2 years ($n = 9$)]: HAQ score at 6 months correlated with baseline DAS28-CRP ($r = 0.705$, $p = 0.02$) and SDAI ($r = 0.678$, $p = 0.03$)
Ramirez García 2014[98]	28	Knees and hands (wrists, MCP, PIP flexor and extensor tendons of the hand)	12 months	Relapse from clinical remission (i.e. a DAS28 of < 2.6)	PDUS	$p = 0.034$; logistic regression model OR 6.18[d]	ESR, TGFβ1	ESR, $p = 0.046$; TGFβ1, $p = 0.082$

Study	Population	Joints	Follow-up	Outcome	Type of US	Correlation of US synovitis with outcome	Type of CE	Correlation of CE with outcome
Saleem 2012[142]	93 RA patients in clinical remission as determined by their treating rheumatologist	Hand and wrist	12 months	Flare – defined as any increase in disease activity requiring a change in therapy [24/93 (26%)]	PDUS and GSUS	PDUS (unadjusted OR 4.08, 95% CI 1.26 to 13.19; $p = 0.014$); GSUS score was not significantly associated ($p = 0.658$)	HAQ-DI, TJC, SJC, RAQoL, CRP, DAS28, SDAI	HAQ-DI per 0.1 unit OR 1.27 (95% CI 1.07 to 1.52; $p = 0.006$). RAQol OR 1.10 (95% CI 1.01 to 1.20; $p = 0.036$). Other variables were NS: SJC, $p = 0.690$; TJC, $p = 0.827$; CRP, $p = 0.308$; DAS28 remission, $p = 0.499$; SDAI remission, $p = 0.616$
Wakefield 2007[143]	10	Bilateral glenohumeral, elbow, wrist, MCP, PIP, knee, tibiotalar, midtarsal and MTP joints	46 weeks	Time to clinical remission	GSUS score PDUS score	$R = -0.221$ (NS) $R = -0.289$ (NS)	Baseline DAS28	$r = 0.627$; $p = 0.071$ (NS trend)
Yoshimi 2014[145]	22 patient with RA in clinical remission (of 31 recruited)	Bilateral wrists and all of the MCP and PIP joints	2 years	Remission, defined by ACR/EULAR Boolean criteria; score of ≤ 1 for all of TJC, SJC, CRP and PGA components	PDUS total score GSUS total score	$p = 0.020$ $p = 0.020$	DAS28-CRP, DAS28-ESR, SJC, TJC, CRP, ESR	NS for all: DAS28-CRP, $p = 0.76$; DAS28-ESR; $p = 0.38$; SJC; $p = 0.060$; TJC, $p = 1.00$; CRP, $p = 0.17$; ESR, $p = 0.39$

CXCL13, chemokine (C-X-C motif) ligand 13; HAQ-DI, Health Assessment Questionnaire Disability Index; NS, non-significant; PGA, patient global assessment; RAQoL, Rheumatoid Arthritis Quality of Life; TGFβ1, transforming growth factor beta 1.

a Multivariable analysis included US joint count, PDUS > 0, DAS, SJC and steroid use.

b Grey-scale USJCAS (ultrasonographic joint count for active synovitis).

c Power Doppler USJIPD (ultrasonographic joint index for power Doppler signal).

d Logistic regression included PDUS, ESR and TGFβ1.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Treatment studies

Treatment response or strategies

Nine studies[79,93,95,100,101,139,154–156] reported data relating to treatment response or strategies. RA was reported as being diagnosed by 2010 ACR/EULAR criteria[11] in two studies[154,155] and by pre-2010 ACR/EULAR criteria[8] in four studies;[100,101,139,156] the criteria used were not reported in two studies.[93,95] Established semiquantitative scoring systems were used by Dougados *et al.*,[139] Inanc *et al.*,[93] Naredo *et al.*[156] (scoring system published by Wakefield[51]), Iwamoto *et al.*[155] (scoring system published by Naredo *et al.*[140]) and Dale *et al.*[154] (scoring system published by Szkudlarek *et al.*[52]). The heterogeneity of the trials precluded meta-analysis. Therefore, no summary estimates of effect are available, which is a limitation of the review. Significance values referred to the association of baseline US and CE measures with treatment measures (such as treatment persistence or response to treatment tapering). These measures could be of importance to patients in so far as they can be used to refine treatment.

Ultrasound was compared with CE as potential predictors of treatment persistence or response (*Table 7*). The study by Ellegaard *et al.*[101] investigated patients starting TNFi treatment to determine whether or not treatment persistence at 1 year of follow-up was associated with US and clinical measures [TJC, SJC, CRP, visual analogue scale (VAS), HAQ and DAS28]. This was a prospective cohort study in which patients were tested with US and underwent CE at baseline and were then followed up and evaluated to investigate the ability of these factors to predict treatment adherence, defined as patients remaining on TNFi therapy at the 1-year follow-up. Among US measures, the square root of the US Doppler colour fraction (USDCF), a measure of hyperaemia, was the only measure that significantly predicted TNFi continuation ($p = 0.008$). None of the clinical measures assessed, including SJC, TJC, DAS28 and CRP, was significantly associated with treatment persistence.[101] In the study by Inanc *et al.*[93], when considering EULAR response compared with no response after 3 months, baseline PDUS ($p = 0.029$) and GSUS ($p = 0.020$) differed significantly between responders and non-responders. This was a prospective cohort study in which patients were tested with US and underwent CE at baseline and were then followed up and evaluated to investigate the association with EULAR response at 3 months' follow-up. Two of the clinical measures assessed also significantly differentiated between responders and non-responders [pain VAS ($p = 0.009$) and SJC ($p = 0.05$)], whereas other clinical measures (DAS28, TJC, ESR and CRP) did not.[93]

Patients with persistent synovitis measured by GSUS or PDUS (*Table 8*), after 4 months of bDMARD treatment, were significantly more likely to have radiological progression at 2 years' follow-up.[139] The study by Dougados *et al.*[139] was a prospective cohort study in which 4 months of biological therapy were prescribed, with further follow-up up to 2 years. The association of lack of response to treatment at 4 months, measured by US or CE using a semiquantitative scoring system, with radiological progression at 2 years was assessed. For persistent synovitis as measured by CE, this did not reach significance.[139] Taylor *et al.*[100] reported prognostic data from a study that was part of a RCT comparing MTX plus IFX with MTX plus placebo in aggressive early RA. Baseline US and CRP were tested for association with radiological progression at 54 weeks. US could significantly predict radiological progression in MTX plus placebo-treated patients ($p = 0.020$), but not in IFX plus MTX-treated patients ($p = 0.479$).[100] However, there were only 12 patients in each group in this study. Baseline CRP did not significantly predict progression in either treatment group.[100]

One prospective cohort study with 6 months' follow-up[155] reported the ability of US and CE to predict relapse following discontinuation of bDMARD (TNFi or TCZ) treatment (*Table 9*). GSUS- and PDUS-detected synovitis at the time of treatment discontinuation (total GSUS score of ≥ 14, total PDUS score of ≥ 3) had higher positive predictive values (PPVs) than DAS28 (≥ 1.5) for predicting relapse within 6 months. Relapse was defined as a DAS28 of > 3.2 in conjunction with escalation of anti-rheumatic treatment. Using these cut-off points patients were more likely to relapse if they had high GSUS ($p = 0.007$) or PDUS ($p = 0.001$) scores, but this was not the case for high DAS28 ($p = 0.297$). Few patients with high US scores did not relapse (GSUS, $n = 2/10$, PDUS, $n = 1/9$), but several patients with high DAS28 did not relapse ($n = 15/28$). However, negative predictive values (NPVs) were similar for GSUS, PDUS and DAS28 as most patients with scores below the cut-off point for US or DAS did not relapse.

TABLE 7 Treatment persistence or response

Study	Population	Joints	Follow-up	Outcome	US type	US association with outcome	CE association with outcome
Ellegaard 2011[101]	109 patients starting TNFis (ADA, ETN or IFX)	Wrist (most affected)	1 year ($n = 69$ with US follow-up)	Treatment persistence (continued with TNFi)	USDCF square root[a]	Treatment persistence vs. dropouts because of lack of efficacy, $p = 0.024$; treatment persistence vs. all dropouts, $p = 0.008$	Treatment persistence vs. dropouts because of lack of efficacy (all NS): TJC, $p = 0.86$; SJC, $p = 0.98$; CRP, $p = 0.86$; VAS (patient global), $p = 0.08$; HAQ, $p = 0.416$; DAS28, $p = 0.943$. Treatment persistence vs. all dropouts (all NS): TJC, $p = 0.321$; SJC, $p = 0.486$; CRP, $p = 0.453$; VAS (patient global), $p = 0.240$; HAQ, $p = 0.098$; DAS28, $p = 0.375$
Inanc 2014[93]	43 patients starting TNFis (drugs NR)	28 joints according to EULAR guideline	3 months	EULAR no response vs. response	PDUS	$p = 0.029$	Pain VAS, $p = 0.009$; SJC, $p = 0.05$; other clinical measures NS (DAS28, $p = 0.90$; TJC, $p = 0.12$; HAQ, $p = 0.31$; ESR, $p = 0.61$; CRP, $p = 0.98$)
					GSUS	$p = 0.020$	
				EULAR good/moderate response (multivariate analysis)[b]	PDUS	OR 0.88 (95% CI 0.68 to 0.94; $p = 0.04$)	NR
					GSUS	NS	

NR, not reported; NS, not significant.
a Other US measures not significant.
b Multivariate analysis included extra-articular involvement, pain VAS and sum score of baseline PDUS.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 8 Progression and treatment response

Study	Population	Follow-up duration	Joints	Outcome	Type of US	US association with outcome	Type of CE	CE association with outcome
Taylor 2004[100]	24 (IFX + MTX, $n = 12$; MTX + PBO, $n = 12$)	54 weeks	Hands and feet	Radiological progression at 54 weeks[a]	Baseline synovial thickness	MTX + PBO, $r = 0.69$, $p = 0.020$; IFX + MTX, $r = -0.23$, $p = 0.479$ (NS)	Baseline CRP	MTX + PBO, $r = 0.58$, $p = 0.077$; IFX + MTX, $r = 0.19$, $p = 0.562$
					Baseline synovial vascularity	MTX + PBO, $r = 0.78$, $p = 0.005$; IFX + MTX, $r = -0.28$, $p = 0.372$ (NS)		
Dougados 2013[139]	59 (ETN, $n = 34$; ADA, $n = 23$; IFX, $n = 2$)	2 years	Wrist, MCP, PIP and MTP joints	Radiological progression[b]	GSUS (> 0 on scale from 0 to 3), persistent synovitis after 4 months of treatment	OR 3.14 (95% CI 1.50 to 6.55; $p = 0.002$)	CE synovitis (> 0 on scale 0–3) persistent synovitis after 4 months of treatment	OR 1.70 (95% CI 0.93 to 3.12; $p = 0.086$)
					PDUS (> 0 on scale from 0 to 3), persistent synovitis after 4 months of treatment	OR 2.79 (95% CI 1.19 to 6.56; $p = 0.019$)		

PBO, placebo.
a vdHSS.
b Occurrence or worsening of erosion or joint space narrowing.

TABLE 9 Accuracy of US to predict relapse following discontinuation of treatment

Study	Population	Follow-up	Joints	Outcome measure	US type[a]	US PPV (95% CI), %	US NPV (95% CI), %	US sensitivity (95% CI), %	US specificity (95% CI), %	CE type[a]	DAS28 PPV, %	DAS28 NPV, %	DAS28 sensitivity, %	DAS28 specificity, %
Iwamoto 2014[155]	40 patients in clinical remission who had discontinued bDMARDs	6 months	134 synovial sites in 40 joints	Relapse DAS28 of > 3.2 and anti-rheumatic treatment escalated	GSUS cut-off point ≥ 14 ($n = 10$)	80 (NR)	73 (NR)	50 (NR)	92 (NR)	DAS28 cut-off point ≥ 1.5 ($n = 28$)	46	75	81	38
					PDUS cut-off point ≥ 3 ($n = 9$)	89 (NR)	74 (NR)	50 (NR)	96 (NR)	DAS28 cut-off point < 1.5 ($n = 12$)	25	NR	NR	NR

NR, not reported.
a Cut-off points determined by receiver operating characteristic (ROC) curve analysis.

This study[155] also compared the median values of several variables for patients with and without relapse (*Table 10*); the total GSUS score differed significantly between the groups ($p = 0.005$), as did the total PDUS score ($p = 0.002$), whereas clinical variables (DAS28, SDAI, CDAI and HAQ) did not. Two other prospective cohort studies[95,156] investigated the association of outcomes with US assessment or CE at the time of treatment discontinuation or tapering (see *Table 10*).

A study of patients in clinical remission with a reducing dose of bDMARDS[156] found that PDUS could significantly predict tapering failure ($p < 0.0005$), defined as an increase in bDMARD dose and/or the presence of clinical disease activity according to both DAS28 and SDAI criteria (see *Table 10*), whereas GSUS was not significantly associated with tapering failure. Some clinical variables could also significantly predict tapering failure, namely DAS28 ($p = 0.011$) and SDAI ($p = 0.003$).[156] In a study of patients in clinical remission discontinuing or tapering DMARDS (bDMARDs or cDMARDs not specified),[95] PDUS significantly predicted disease flare ($p = 0.06$ by multivariate analysis) (see *Table 10*), as did the number of DMARDs taken at baseline, whereas DAS28 did not.

One RCT[154] investigated treatment strategies with and without US (*Table 11*). In this study, patients were given the same treatment for 3 months and then either step-up DMARD escalation strategies guided by DAS28 alone ($n = 57$) (target = DAS28 of < 3.2) or step-up DMARD escalation strategies guided by DAS28 plus PDUS assessment of a limited joint set ($n = 53$) (target = PDUS signal in one or fewer joints). Dale *et al.*[154] found that, at 18 months, significantly more patients in the group receiving PDUS had attained Disease Activity Score 44 joints (DAS44) remission ($p = 0.046$). Since drafting this report, the full results of this study have been published.[78] At 18 months' follow-up, significantly more patients in the group receiving PDUS had attained DAS44 remission ($p = 0.03$) (see *Table 11*). However, for the two co-primary end points, change from baseline in DAS44 and rheumatoid arthritis magnetic resonance imaging scoring system (RAMRIS) erosion score, there was no significant between-group difference ($p = 0.72$ and $p = 0.33$, respectively). A preliminary publication of the ARCTIC study results,[79] identified post search, also investigated treatment strategies with and without US in a randomised comparison (see *Table 11*). The primary end point (a DAS of < 1.6 and no swollen joints at 16, 20 and 24 months and no progression in vdHSS between 16 and 24 months) was reached by 26 patients in the US strategy group and 21 in the control group, with no significant difference between the groups.

For both randomised trials (TaSER and ARCTIC), there was no significant difference in primary end point between the clinical strategy group and the strategy group based on US and clinical measures. Both trials found a significant advantage of the strategy including US for one of the secondary outcomes (DAS remission in the TaSER study and erosion score in the ARCTIC study), but not for the other secondary outcomes (ACR core set variables in the TaSER study; CDAI or SDAI remission or EULAR response in the ARCTIC study). This suggests that adding US to a DAS-based strategy is not beneficial after 18–24 months. It is uncertain whether this is because of a genuine lack of improvement caused by the US strategy or because of other factors. Both trials included early RA patients only and PDUS (the ARCTIC study also used GSUS) and have some incorporation bias in terms of the clinical strategy and primary outcome being DAS related (although in the TaSER study the strategy used DAS28 and one of the co-primary outcomes was DAS44 and in the ARCTIC study the primary outcome was a composite measure of remission that included DAS).

Treatment decisions

Six studies[89–91,94,152,153] reported on the use of US in addition to clinical measures and the impact on treatment decisions (*Table 12*). These were fairly small studies, with sample sizes ranging from 17[89] to 109[94] (see *Appendix 8*). Four studies[89,90,94,153] took place in rheumatology clinics in the UK, one[152] was from the USA and one[91] was from France. Four studies[89–91,152] did not report the source of funding, although three[89–91] of these stated that the authors had no conflicts of interest. Two studies[94,153] were supported by pharmaceutical companies. Dale *et al.*[153] stated that neither of their funders had any role in the design, performance, analysis, interpretation or reporting of the study.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 10 Tapering and treatment discontinuation

Study	Population	Follow-up	Joints	Outcome	US type	US association with outcome	CE association with outcome
Iwamoto 2014[155]	42 RA patients in clinical remission, discontinued biological therapy[a]	6 months	134 synovial sites in 40 joints	Relapse (i.e. a DAS28 of > 3.2) and anti-rheumatic treatment escalated	Total GSUS score Total PDUS score	$p = 0.005$ $p = 0.002$	All NS: DAS28-ESR, $p = 0.609$; DAS28-CRP, $p = 0.389$; SDAI, $p = 0.180$; CDAI, $p = 0.275$; HAQ-DI, $p = 0.721$
Luengroongroj 2015[95]	32 RA patients in clinical remission on DMARD(s), about to stop or reduce dose[b]	3 months	NR	Disease flare (arthritis symptoms and signs detected)	PDUS (median total score)	Univariate analysis: OR 2.14 (95% CI 1.13 to 4.05) (significant; *p*-value NR). Multivariate analysis:[c] NS (trend); OR 3.06, (95% CI 0.95 to 9.84; $p = 0.06$)	Number of DMARDs: univariate analysis – OR 5.88 (95% CI 1.12 to 30.88) (significant); in multivariate analysis NS. DAS28-CRP: NS by univariate or multivariate analysis
Naredo 2015[156,157]	77 RA patients in sustained clinical remission[d]	12 months ($n = 35$, 45.5%, tapering failure at 12 months)	42 (including hands and feet)	bDMARD tapering failure (increase in bDMARD doses and/or the presence of clinical disease activity according to both DAS28 and SDAI criteria)	PDUS global index for Doppler synovitis (DSI) calculated GSUS	Significant associations: higher DSI ($p < 0.0005$); and DSI > 0 ($p < 0.0005$) at baseline. Multivariate analysis: DSI > 0 (OR 29.92, 95% CI 6.81 to 131.40; $p < 0.0005$) GSUS was not a significant predictor of tapering failure in the multiple regression analysis[e]	Significant associations: longer duration of RA ($p = 0.009$), higher number of previous synthetic DMARDs ($p = 0.003$), higher DAS28 ($p = 0.011$), higher SDAI ($p = 0.003$) at baseline. Multivariate analysis: DAS28 of ≥ 2.2 (OR 5.81, 95% CI 1.62 to 20.93; $p = 0.007$). Other clinical variables (duration of remission) were not significant predictors in multiple regression analysis

NR, not reported, NS, not significant.
a Remission: DAS28 of < 2.6.
b Remission: CDAI of < 2.8.
c Multivariate analysis included number of DMARDs, anti-cyclic citrullinated peptide antibody level, mean DAS28-CRP and median total PDUS score.
d A DAS28 of < 2.6 or a SDAI of < 3.3.
e Multivariate analysis included baseline DAS28 and the global index for Doppler synovitis (DSI) for 42 joints, 12 joints and wrist–MCP–ankle–MTP.

TABLE 11 Treatment strategies with and without US

Study	Population	Follow-up duration	Outcome	US treatment strategy group	Clinical treatment strategy group	Comparison
Dale 2013[154] (TaSER)	110 early RA patients randomised to 3 months of the same treatment and then step-up DMARD escalation strategies guided by either DAS28 alone ($n = 57$) (target: DAS28 of < 3.2) or DAS28 + PDUS assessment of a limited joint set ($n = 53$) (target: PDUS signal in one or fewer joints)	18 months	DAS44 remission (i.e. a DAS44 of < 1.6)	Strategy PDUS and DAS28: significant improvement in DAS44 (mean change −2.76, 95% CI −0.84 to 0.33; $p = 0.39$). DAS44 remission: 33% at 3 months, 65% at 18 months	Strategy DAS28: significant improvement in DAS44 (mean change in DAS44 −2.51, 95% CI NR). DAS44 remission: 43% at 3 months, 44% at 18 months	After 18 months, more patients in the PDUS group had attained DAS44 remission ($p = 0.046$) (at 3 months, difference was not significant: $p = 0.32$)
Dale 2016[78] (TaSER)	111 RA or undifferentiated arthritis patients[a]	18 months	DAS44 remission (i.e. a DAS44 of < 1.6)	($n = 54$) DAS44 remission: 35 (66%)	($n = 57$) DAS44 remission: 25 (43%)	DAS44 remission ($p = 0.03$)
			Mean change in DAS44	–2.69 (significant improvement)	–2.58 (significant improvement)	Non-significant difference between groups ($p = 0.72$)
			Median (IQR) change in RAMRIS erosion score	0.5 (0.0–1.0)	1.0 (0.0–2.0)	Non-significant difference between groups ($p = 0.33$)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 11 Treatment strategies with and without US (*continued*)

Study	Population	Follow-up duration	Outcome	US treatment strategy group	Clinical treatment strategy group	Comparison
Haavardsholm 2015[79] (ARCTIC)	130 early RA patients	24 months	Primary end point was a DAS of < 1.6, a SJC44 of < 1 and a ΔvdHSS of < 0.5 between 16 and 24 months	(*n* = 118) Strategy PDUS and GSUS and DAS: 26 (22.0%)	(*n* = 112) Strategy DAS: 21 (18.8%)	*p* = 0.54
			DAS remission (i.e. a DAS of < 1.6)	80 (67.8%)	75 (67.0%)	*p* = 0.89
			Median (IQR) change in vdHSS total score between 0 and 24 months	1.0 (0–2.5)	1.5 (0.5–3.0)	*p* = 0.09
			Median (IQR) change in vdHSS erosion score between 0 and 24 months	0.5 (0–1.5)	1.0 (0.5–2.0)	*p* = 0.04

IQR, interquartile range; NR, not reported.
a Only one patient did not meet 2010 ACR/EULAR RA classification criteria.[11]

TABLE 12 Use of US in addition to CE alone and impact on treatment decisions

Study	Source of funding	Population, number of patients	Setting	US and treatment decisions
Bhamra 2014[89]	NR	17	Nuffield Orthopaedic Hospital Emergency Rheumatology Clinic, UK	In 15/17 (88%) patients, treatment escalation was directly influenced by GSUS findings
Ceponis 2014[152]	NR	51	San Diego Health System, CA, USA	US use modified the bDMARD and/or cDMARD used in seven cases, but did not affect the overall treatment plan ($p > 0.15$) or overall cDMARD ($p < 0.062$) or bDMARD use ($p > 1.0$). Use of US increased physicians' confidence ($p < 0.0005$) and patients reported that their confidence in physicians' medical decisions had increased (88.4% of cases)
Ciurtin 2013[90]	NR	39	Department of Rheumatology, University College Hospital, London, UK	Study identified 9/39 (23%) RA patients with active synovitis with a positive Doppler signal that prompted a change of treatment
Dale 2014[153,161] (TaSER)	Chief Scientist's Office, Scottish Executive, and Pfizer UK	53	Three Glasgow teaching hospital sites, UK	On 120 occasions (29%), GSUS findings contradicted the DAS28 and led to modified treatment decisions
Gandjbakhch 2008[91]	NR	52	University rheumatology centre, Paris, France	US results caused a change in treatment in 13% of patients. Confidence of the clinician, measured using a VAS (0–100), increased by 11 points (95% CI 5.9 to 16.9 points) following US ($p < 0.001$)
Kelly 2013[94]	AbbVie	109	Four secondary care rheumatology clinics: two in London, one in Southampton and one in Antrim, UK	The US-monitored group ($n = 54$) had a significantly shorter time to initiation of DMARDs (1.45 vs. 2.38 months) than the non-US group ($n = 55$) (t-test, $p = 0.02$)

NR, not reported.

When the percentage of treatment decisions modified by the additional use of US was reported, this ranged from 13% to 88% of cases[89–91,153] (for studies based in the UK the percentages were 23%,[90] 29%[153] and 88%[89]).

One study[89] examined treatment decisions in clinician-evaluated patients by survey and found that 15 out of 17 (88%) treatment escalations were influenced by the addition of US. Of two observational studies, one[91] found that US (in addition to CE and DAS28) influenced treatment decisions in only 7 out of 52 (13%) patients; however, clinician confidence was significantly improved following US ($p < 0.001$). The other observational study[94] found a significant difference in time to initiation of DMARDs ($p = 0.02$) (the US-monitored group had a shorter time to initiation than the group undergoing clinical evaluation alone) (see *Table 12*). Two studies investigated the utility of routine US and reported a change in treatment following US in 9 out of 39 (23%) patients with clinician-evaluated synovitis[90] and a modification of the bDMARD or cDMARD used in 7 out of 51 patients (compared with clinical evaluation of swollen joints and CDAI);[152] clinician confidence was also significantly improved following US ($p < 0.0005$) in one of the studies.[152] A study of treatment strategy[153] reported that, on 120 occasions (29%), GSUS findings contradicted the DAS28 and led to modified treatment decisions (see *Table 12*).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Survey

A survey of UK rheumatology units (see *Appendix 1*) with 31 respondents suggested that US is already being used in some units for modifying treatment decisions in RA. Twenty respondents (65%) said that they used US for monitoring synovitis. Additionally, one respondent said that their unit planned to use US in the future. Survey respondents were self-selecting and so the sample may have been biased. Only 31 responses were received by the end of February 2016 and the small sample size is a limitation of the survey.

Discussion

This systematic review aimed to investigate the value of US in addition to CE for monitoring synovitis and whether or not US could be used to guide treatment decisions. There were few RCTs available and thus lower-quality study designs were included.

Thirty-three[53,92,96,102–116,118–132] studies provided diagnostic data (see *Appendix 7*). The majority of these studies reported that US detected more cases of synovitis than CE alone. The detection of subclinical synovitis would be useful only if clinically relevant, with prognostic studies suggesting that US-detected synovitis was associated with radiographic progression.

Power Doppler ultrasound was significantly associated with radiographic progression in all studies in which it was measured. GSUS was significantly associated with radiographic progression in some, but not all, studies in which it was measured. Similarly, DAS was significantly associated with radiographic progression in some, but not all, studies in which it was measured. Studies varied in terms of interventions, comparators and outcomes, making it difficult to draw firm conclusions from the available evidence. Few studies were identified that compared US with CE and their effect on treatment; however, these studies suggested that US was superior to CE alone in predicting response to treatment tapering or discontinuation. Few data were identified regarding the additional influence of US in current practice, but studies suggested that US was used in treatment decisions and could increase physician confidence in those decisions. A small survey of UK rheumatology units (see *Appendix 1*) suggested that US is already being used in some units for modifying treatment decisions in RA.

US could also distinguish synovitis from inflammation resulting from other pathologies. Most of these studies assessed synovitis in hand and wrist joints, reporting that US detected more cases of synovitis in these joints than CE; this was also the case for elbow and shoulder joints. Foot and ankle joints were less likely to show an advantage of US over CE. The majority of studies investigating responsiveness found similar responsiveness for US and clinical comparator measures including DAS28[110,111,113,116,132] SDAI[111,116,117] CDAI[111,113] SJC[111] TJC[111] and CRP.[129]

Fifteen studies[55,69,97,98,113,134,135,138–144,147] provided prognostic data. Although the design of these studies was of high quality (prospective cohort), data reported were correlations and sample sizes ranged from 10 to 453 (in total, data were available from 1523 patients but study heterogeneity precluded meta-analysis). The majority of studies investigating radiographic progression reported that US was a significant predictor, either GSUS[55,69,134,135,138,139] or PDUS.[55,69,97,113,135,138–140,144,160] There were mixed results regarding the association of clinical comparators with outcome measures. Regarding other outcome measures, PDUS was a significant predictor in the majority of studies for DAS28,[55,98,137] ACR/EULAR remission[145] and flare.[142] There were conflicting results for the association between HAQ score and US.[55,97] The data suggested that there was a stronger association with outcomes for PDUS than for GSUS.

Nine studies[79,93,95,100,101,139,154–156] reported data relating to the impact of US on treatment response or strategies. Two were RCTs[79,154] and the others were prospective cohort studies [93,95,100,101,139,155,156] Sample sizes ranged from 24 to 130, with data available from 627 patients in total (although study heterogeneity precluded meta-analysis). Six small (sample size 17–109 patients; 321 patients in total) studies[89–91,94,152,153] reported observational data on the impact of US on treatment decisions.

One study[101] found that US was the only measure that significantly predicted TNFi continuation. One study[93] found that baseline PDUS and GSUS differed significantly between EULAR responders and non-responders. US-measured treatment response was associated with radiological progression[100,139] and PDUS and GSUS were better predictors of relapse following treatment discontinuation than DAS28.[155] PDUS was more highly associated with tapering failure than clinical variables[156] and PDUS significantly predicted disease flare on treatment whereas DAS28 did not.[95]

Randomised controlled trial evidence of early RA patients reported that the addition of PDUS to a DAS28-based treat-to-target strategy led to significantly more patients attaining DAS44 remission, but there was no significant between-group difference in change from baseline in DAS44 and RAMRIS erosion score.[154] In the ARTIC study,[162] there was no significant between-group difference (PDUS, GSUS and DAS vs. DAS) in the primary end point, which consisted of a composite of a DAS of < 1.6 and no swollen joints at 16, 20 and 24 months and no progression in vdHSS between 16 and 24 months. Both RCTs were subject to incorporation bias, with DAS being used in the treatment strategy groups and forming part of the primary outcome measure; however, the use of DAS reflects real-world practice. Outcome measures not subject to incorporation bias were the erosion scores, with RAMRIS not being significantly different between groups in the TaSER study[154] and the vdHSS erosion score having a significant advantage for the strategy with additional US in the ARCTIC study.[162] Only two RCTs[154,162] were found comparing US strategies with clinical strategies alone. In clinical practice, a diagnostic test should only be considered when there is clinical uncertainty. These two studies did not explore the value of US when added only in cases of clinical uncertainty, for example, when there was discrepancy between DAS and clinical evaluation. Several small studies have used US in combination with clinical measures for treatment decisions.[26,163,164] Most studies relied on one-off US measurements at baseline; however, the two RCTs employed US serially as part of treatment strategies. Heterogeneity of trials precluded meta-analysis. Therefore, no summary estimates of effect are available, which is a limitation of the review.

Treatment decisions were modified by the additional use of US in 13%–88% of case.[89–91,153] Clinician confidence was significantly improved following US.[91,152] A study of treatment strategy[153] reported that, on 120 occasions (29%), GSUS findings contradicted the DAS28 and led to modified treatment decisions.

Guidelines and reviews were identified that were of relevance to this report,[9,24,25,28–30,45,46,50,75–77,85–88] although none had a scope that was identical to the systematic review conducted here. The bibliographies of these guidelines and reviews were searched and studies meeting the systematic review inclusion criteria were included in the review.

Guidelines and reviews differed in scope from this report by not being restricted to RA[29,46,85–87] or by looking at the initial diagnosis of RA,[28,30,46,88] by not being restricted to synovitis[9,25,28–30,46,75–77] or not including synovitis[85] or by focusing on management[9,76,77] or not covering management decisions.[24,25,86,88] Some imaging reviews differed from the review in this report by including US-guided injections,[30,46] by not being restricted to US as an imaging technique[28,29,45,87] or by focusing on inter-rater reliability.[50]

The EULAR guidelines on imaging in RA[28] covered all of the questions addressed by the scope of this report, apart from the percentage of treatment decisions influenced by the addition of US to CE. However, as these guidelines were published in 2013, they could not include the more recent publications that this report included. As the studies relating to treatment identified in this report were nearly all from 2013 or later, only one[101] of these studies was included in the published EULAR guidelines.[28]

The findings of this report were in agreement with these published guidelines and reviews in that US detects synovitis that is not apparent on CE and that US detects synovitis in patients in clinical remission.[24,25,28,30,45,88] The findings of this report were also in agreement with these published guidelines and reviews in that US can assess the course of disease,[30,76,87] with similar responsiveness to that of DAS28.[28]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

With regard to prognosis, the findings from this report agreed with the published guidelines and reviews in that PDUS is a predictor of erosive progression and disease flare.[25,28,76,88]

Other reviews have suggested that there is still a need for more evidence regarding which joints should be assessed by US for the most effective practice.[24,45]

Most studies were carried out outside the UK and, as such, treatment pathways may differ from treatment pathways in UK patients. However, the results from this report are generalisable to UK practice in terms of studies using international definitions of RA. Additionally, most used standardised semiquantitative scoring systems for US measures of synovitis and clinical measures that are used internationally, for example DAS28. Treatments assessed included cDMARDs and bDMARDs, as would be relevant to UK practice.

Chapter 4 Assessment of cost-effectiveness

Literature reviews undertaken

Reviews were undertaken to identify literature on the:

- cost-effectiveness of the use of US for monitoring synovitis in RA
- economic impact of tapering bDMARDs in the treatment of RA.

For brevity these have been denoted the 'monitoring' search and the 'tapering' search, respectively. The search strategies for these searches are provided in *Appendix 3*.

Following deduplication, 226 hits were obtained in the monitoring search and 54 hits were obtained in the tapering search. Abstracts were sifted by two reviewers and, because of the relatively small numbers of hits obtained, a sensitive approach was taken in which any paper felt to be of potential relevance by either reviewer was retrieved as a full paper. From the sift, five papers were retrieved from the monitoring search[165–169] and 19 papers were retrieved from the tapering search.[170–187] All papers were reviewed and, when their reference lists indicated that further papers would potentially be of benefit, these were also reviewed.

Papers potentially relating to the cost-effectiveness of ultrasound for monitoring synovitis or tapering drug doses

Three cost-effectiveness studies relevant to the decision problem were identified although all of these were related to tapering of the amount of drugs used.[176–178] The first study by Kobelt *et al.*[176] concluded that, in a situation in which a considerable proportion of patients achieve remission, dose adjustments would increase the cost-effectiveness of ETN. A similar conclusion was produced in a later publication by Kobelt *et al.*,[177] whereby dose reduction was assumed to be the most cost-effective strategy for patients with moderate disease rather than full-dose ETN or no ETN use. Krieckaert *et al.*[178] assessed the cost-effectiveness of personalised treatment for RA based on clinical response and drug levels at 6 months. For patients who had a EULAR response at 6 months, a decision to stop treatment, continue treatment or prolong the interval between ADA doses was made based on the drug levels detectable within patients. The authors concluded that tailoring ADA based on short-term outcome and drug levels was cost-effective compared with usual care.

No papers were identified on the cost-effectiveness of US use in monitoring levels of synovitis. Thus, any analysis undertaken within this report is, to the authors' knowledge, the first such assessment. The consensus from the three papers regarding the cost-effectiveness of tapering RA treatment indicates that dose reductions would represent a cost saving that could be used in other areas of the health service; in addition, providing that response and/or remission was maintained, dose reductions would also be beneficial to patients because of the avoidance of potential adverse events associated with taking bDMARDs. A review of adverse events associated with bDMARDs within RCTs is reported by Stevenson *et al.*[23]

Papers potentially relating to the efficacy of conventional or biological disease-modifying anti-rheumatic drugs when the dose has been tapered

Evidence regarding the continued efficacy of cDMARDs or bDMARDs following a reduction in dose or withdrawal of treatment was identified in papers retrieved in the literature review (both clinical and cost-effectiveness reviews) and from references within these papers. When abstracts were identified,

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

searches for a later publication were undertaken. Only the most recent papers were included when multiple papers related to the same study unless the older papers contained data that were not included in the more recent paper. Furthermore, papers known to our clinical experts but not identified by other means were also included.

A limitation of this approach is that the sensitivity of the search strategies was not high, with a minority of the final set of papers also identified within the tapering search. As such, potentially relevant papers may have been omitted, although for the purposes of this report this was deemed acceptable as the intention was to provide a broad overview rather than to fully detail the evidence.

A brief summary of the methods and conclusions from these papers is provided in *Table 13*, with papers listed in reverse chronological order. The evidence suggests that drug tapering, treatment holidays or treatment withdrawal in patients induced into remission or low disease activity can occur without harm in a sizeable proportion of patients, although the evidence for withdrawal of treatment in patients receiving ADA or CTZ is arguably weaker than that for IFX. Tapering drug doses would have many benefits including reduced drug acquisition costs and potential avoidance of the adverse events associated with DMARDs.

TABLE 13 Results from a review of dose tapering studies within RA

Study	Setting	Summarised details[a]	Summarised authors' conclusions
Fautrel 2016[188]	France	An 18-month equivalence randomised trial including (1) patients receiving ETN and ADA at stable doses for ≥ 1 year, (2) patients in remission according to the DAS28 for ≥ 6 months and (3) patients with stable joint damage. In total, 137 patients were randomised to standard maintenance ($n = 73$) or to a regime using an algorithm to increase the period between injections based on the DAS28, every 3 months, until completely stopping ($n = 64$). In the algorithm arm bDMARDs were stopped for 39%, tapered for 36% and maintained at full dose for 20%. The status of the remaining patients (5%) was not clear. Relapse was more common in the algorithm arm (76.6% vs. 46.5%; $p = 0.0004$). However, there was no difference in structural damage progression ($p = 0.8$)	The tapering algorithm was not equivalent to the maintenance strategy, resulting in more relapses without impacting structural damage progression
Galloway 2015[189]	UK	A pragmatic 12-month RCT evaluating if tapering bDMARDs causes loss of response. In total, 103 patients receiving ETN or ADA and with a DAS28 of < 3.2 for > 3 months were randomised to one of three groups for 6 months: (1) constant bDMARD dose, (2) 33% tapered dose or (3) 66% tapered dose. Flares, defined as an increase in DAS28 to ≥ 3.2 and one or more swollen joint, occurred in 7/50 (14%) control subjects, 6/48 (13%) in the 33% tapering group and 14/38 (37%) in the 66% tapering group. Post-tapering flares resolved when TNFi was restarted. The OR for a flare in the 33% tapering group compared with the 66% tapered group was 4.2 (95% CI 1.3 to 14.5). There were no significant differences at 6 months (*p*-value not reported)	Good responses are maintained after bDMARD doses are tapered by one-third. Tapering by two-thirds results in more flares, but these respond to restarting bDMARDs and did not adversely affect disability progression. The 33% tapering strategy retains responses at substantially reduced drug costs

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
Ghiti Moghadam 2015[190]	The Netherlands	An open-label RCT assessing whether RA patients with a DAS28 of < 3.2 in the previous 6 months can effectively and safely stop and restart bDMARDs. Patients were randomised 2 : 1 to stop or continue their current bDMARD. A flare was defined as a DAS28 of ≥ 3.2 with an increase of ≥ 0.6 compared with the previous DAS28. In total, 531 patients were randomised to stop treatment and 286 were randomised to continue treatment. At 6 months, significantly more patients in the stop group (29.3%) had experienced a DAS28 flare than in the continuation group (9.7%) ($p < 0.0001$). At 12 months, these values were 40.8% and 14.4%, respectively ($p < 0.0001$). Of the patients who restarted a TNFi within the first 26 weeks after stopping, 83% had regained a DAS28 of < 3.2 6 months later, with a median time to regained DAS28 of < 3.2 of 12 weeks (95% CI 10.9 to 13.1 weeks)	During a 12-month follow-up period, 59% of RA-patients with a DAS28 of < 3.2 were able to stop bDMARDs without experiencing a flare. The data suggest that bDMARDs can be restarted effectively and safely
Luengroongroj 2015[95]	Thailand	A prospective study to assess whether or not US can be used to predict relapse after tapering of DMARDs. Thirty-two patients with established RA and clinical remission defined by a CDAI of < 2.8 were enrolled. Patients on a maintenance dose had their dose reduced to a half-dose, whereas those on lower doses had their treatment withdrawn. A relapse was defined as having arthritis symptoms and clinical signs. Four patients (12.5%) relapsed within 3 months	Reducing doses of DMARD for a short period of time appears to be safe; however, close monitoring for disease relapse is needed, especially in patients with subclinical synovitis
Marks 2015[26]	England	Prospective cohort study analysing the possibility of reducing the dose of bDMARDs by one-third. Seventy patients were recruited who were in disease remission (i.e. a DAS28 of ≤ 2.6), had had an absence of synovitis on PDUS for > 6 months and who were not taking corticosteroids. Combined DAS28 and PDUS remission was maintained by 96% at 3 months, 63% at 6 months, 37% at 9 months and 34% at 18 months. Those who continued on the reduced doses were more likely to have lower DAS28 scores on initiation of bDMARD therapy and to be rheumatoid factor negative	Combined clinical and US assessment identifies individuals in remission who may be suitable for bDMARD dose reduction
Tanaka 2015[191]	Japan	Prospective cohort study analysing the possibility of discontinuing ADA for 1 year. A flare was defined as a DAS28 of ≥ 3.2. In total, 75 patients who were steroid free and with a DAS28 of < 2.6 for 6 months participated in the study; 52 patients chose to discontinue ADA whereas 23 chose to remain on	The possibility of remaining ADA free for a year was demonstrated, particularly in those with deep remission

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
		treatment. The proportion of patients maintaining remission was significantly higher in the continuation group. In patients with deep remission (i.e. a DAS28 of ≤ 1.98) there was no significant difference between the groups. ADA readministration to those with a flare was effective	
van Herwaarden 2015[187]	The Netherlands	A randomised, controlled, open-label, non-inferiority study ($n = 180$) evaluating disease activity-guided dose reductions of ADA or ETN ($n = 121$) vs. usual care ($n = 59$) over an 18-month period. Patients in the dose reduction arm increased the injection interval every 3 months until flare, defined as a DAS28 score increase of > 1.2 or an increase of > 0.6 and a DAS score of ≥ 3.2, or bDMARD discontinuation. Following a flare the last effective dosing regimen was reinstated. A major flare was defined as a flare with a duration of > 3 months. Dose reduction was non-inferior to usual care: 12% had a major flare in the dose reduction arm compared with 10% in the usual care arm. bDMARDs could be successfully stopped in 20% of patients, but dose reduction was reportedly not possible in 37%. Although functional status, quality of life, relevant radiographic progression and adverse events did not differ between strategies, flares (73% vs. 27%) and radiographic progression (32% vs. 15%) were more frequent in the dose reduction arm	A strategy of disease activity-guided dose reduction is non-inferior to usual care with regard to major flaring, resulting in successful dose reduction or stopping in two-thirds of patients
van Vollenhoven 2015[192]	Europe (16 European sites)	A randomised, double-blind trial assessing the impact of dose reductions of ETN in 91 patients with stable low disease activity (i.e. a DAS28 of ≤ 3.2) receiving ETN treatment. Following an 8-week screening period, 73 patients were randomised to the usual dose of ETN (50 mg) ($n = 23$), a half-dose (25 mg) ($n = 27$) or placebo ($n = 23$). Sixty-six patients completed the study. The proportions of patients who did not flare at week 48 were 52% for full-dose ETN, 44% for half-dose ETN and 13% for placebo. The majority of patients who flared regained low disease activity with 50 mg of ETN	In patients who have achieved stable low disease activity on ETN, continuing treatment is superior to placebo. Reduced-dose ETN was also more effective than placebo in maintaining a favourable response, suggesting that a maintenance strategy with reduced-dose ETN may be possible
Emery 2014[193]	Global (57 sites in Europe and Asia)	In total, 306 patients with early active disease were enrolled and treated with ETN plus MTX. A total of 193 patients with a DAS28 of ≤ 3.2 at week 39 of treatment and a DAS28 of < 2.6 at week 52 were randomised to receive one of 25 mg of ETN plus MTX, MTX alone or placebo for 39 weeks. Patients with a DAS28 of ≤ 3.2 at week 39 of the randomised period had all study	Patients with early RA who achieved remission while receiving 50 mg of ETN plus MTX had better disease control with 25 mg of ETN plus MTX than MTX alone or placebo. No significant difference was observed in radiographic progression between the three groups

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
		treatment withdrawn. At the end of the randomised period, significantly more patients on combination therapy had sustained remission than patients receiving the remaining strategies: 40/63 (63%) of the combination therapy group, 26/65 (40%) of the MTX group and 15/65 (23%) of the placebo group. Following the treatment withdrawal period, the combination therapy group was no longer significantly better at sustaining remission than the MTX group ($p = 0.10$): 28/63 (44%) of the combination therapy group, 19/65 (29%) of the MTX group and 15/65 (23%) of the placebo group. No significant between-group differences were observed in radiographic progression of disease ($p \geq 0.34$ for all comparisons). Serious adverse events were reported in three patients (5%) in the combination therapy group, two (3%) in the MTX alone group and two (3%) in the placebo group (*p*-value not reported)	
Iwamoto 2014[155]	Japan	A prospective study of patients on bDMARDs to assess whether or not US assessment of synovitis predicts relapse after withdrawal of bDMARDs. 42 patients were enrolled who were in clinical remission (i.e. a DAS28 of < 2.6) and who agreed to withdraw treatment. The mean duration of remission was 23 months (range 3–73 months)	Comprehensive US assessment predicted relapse within a short term after discontinuation of bDMARDs
Naredo 2014[156]	Spain	77 patients treated with bDMARDs were recruited if they met the following criteria: (1) stable dose of treatment in the previous 12 months, (2) sustained clinical remission based on DAS28 or SDAI in the previous 12 months, (3) ≤ 5 mg/day of prednisone treatment in the previous 6 months and (4) not having needed NSAIDs for > 1 week nor local corticosteroid injections in the previous 6 months. bDMARDs were tapered by increasing the duration between doses or reducing the dose. Tapering failure (assessed at 6 and 13 months) was defined as an increase in the dose and/or disease activity on the DASA28 or SDAI. In total, 23 (29.9%) patients were tapering failures at ≤ 6 months and 35 (45.5%) were tapering failures at 12 months. Significant predictors of failure at 6 months ($p \leq 0.05$) were longer duration of RA, a higher number of synthetic DMARDs, a higher DAS28 at baseline, a higher SDAI at baseline, a higher global index for Doppler synovitis (DSI) at baseline and a DSI > 0 at baseline	The results suggested that the presence and grade of Doppler-detected synovitis may predict biologic therapy tapering failure in RA patients in sustained clinical remission

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
Smolen 2014[182]	Global (161 sites worldwide)	Patients with low disease activity [i.e. a DAS28 of < 3.2 at weeks 22 and 26 from the Optimal Protocol for Treatment Initiation with Methotrexate and Adalimumab (OPTIMA trial)[194]] were randomised to either ADA continuation or withdrawal for a period of 52 weeks or, if not on ADA, were maintained on MTX monotherapy. A total of 207 patients on ADA were rerandomised (105 continued to take ADA); 112 patients continued on MTX monotherapy. In total, 95/103 (92%) patients continuing ADA treatment maintained a DAS28 of < 3.2 compared with 75/90 (83%) whose treatment was withdrawn ($p = 0.0569$); 90/103 (87%) continuing ADA had a DAS of < 2.6 compared with 62/90 (69%) whose treatment was withdrawn ($p = 0.017$)	Outcomes were similar regardless of whether ADA was continued or withdrawn in patients who initially responded to ADA
Aguilar-Lozano 2013[195]	Mexico	A prospective cohort study of patients ($n = 45$) in remission (i.e. a DAS28 of ≤ 2.6) with no swollen joints following cessation of TCZ. In total, 20 patients maintained remission during the 12-month follow-up period. Of the 25 who relapsed, 14 (56%) did so within 3 months of the last dose of TCZ	Long-term clinical remission is possible in a number of patients with RA after the suspension of TCZ
Detert 2013[196]	Germany	Prospective trial comparing the use of ADA with placebo for 24 weeks after which both groups received MTX alone for 24 weeks. The trial recruited 172 people with active early (≤ 12 months) RA. At week 48, there was no statistically significant difference in DAS28 score between the groups ($p = 0.41$), although changes in radiographic progression significantly favoured the ADA group	A greater, and significant, reduction in radiographic progression was seen in the ADA arm, but this was not the case for DAS28
Holroyd 2013[163]	England	A prospective study of people in clinical and US remission (i.e. a DAS28 of < 2.6 and PDUS = 0) who had their dose of bDMARDs reduced by one-third. In total, 56 of 321 patients met the eligibility criteria. Of these, 42 (75%) remained on the tapered dose for a mean of 8.7 months. Fourteen patients (25%) flared and returned to the full dosages at a mean of 5.9 months	Using US alongside clinical assessment may increase the likelihood of selecting patients who could successfully reduce the dose of bDMARDs while maintaining clinical and US remission
Nishimoto 2013[27]	Japan	Tis study examined the possibility of drug-free remission induced by TCZ monotherapy. In total, 187 patients were enrolled who had achieved a DAS28 of < 3.2. Loss of efficacy was defined as either a DAS28 of > 3.2 at two consecutive visits or initiation of additional treatments on patient request. A DAS28 of < 3.2 was maintained in 65 patients at 24 weeks after discontinuation and 24 patients at	TCZ monotherapy may induce bDMARD-free remission or low disease activity without concomitant use of synthetic DMARDs

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
		week 52; 19 patients (10%) were drug free for 52 weeks, with 17 patients meeting the criteria for remission (i.e. a DAS28 of < 2.6). Low serum interleukin 6 (IL-6) and normalisation of matrix metalloproteinase 3 (MMP-3) levels at cessation of TCZ monotherapy were identified as independent predictive markers for longer duration of low disease activity	
Smolen 2013[183]	Global (80 sites in Europe, Latin America, Asia and Australia)	RCT (PRESERVE) of adult patients with an initial DAS28 of > 3.2 and ≤ 5.1 that had been reduced to low disease activity (mean DAS28 of ≤ 3.2 between weeks 12 and 36 and a DAS28 of ≤ 3.2 at week 36). Patients were assigned to 50 mg of ETN ($n = 202$), 25 mg of ETN ($n = 202$) or placebo ($n = 200$). In total, 166/201 (83%) of those receiving 50 mg of ETN had low disease activity compared with 84/197 (43%) of those receiving placebo and 159/201 (79%) of those receiving 25 mg of ETN. Significantly more patients receiving either dose of ETN had low disease activity than patients receiving placebo ($p < 0.0001$). The authors reported that there was no significant difference between the two ETN doses ($p < 0.379$)	Conventional or reduced doses of ETN maintain low disease activity more effectively than placebo
Smolen 2014[197]	Global (31 European sites)	52-week double-blind RCT (CERTAIN) including a 24-week treatment period and 28-week follow-up period in patients with low to moderate disease activity and stopping therapy in patients in sustained remission. Patients had a CDAI of > 6 and ≤ 16, two or more tender joints, two or more swollen joints and either a ESR of ≥ 28 mm/hour or CRP level of > 10 mg/l at screening and baseline. In total, 194 patients were randomised to placebo ($n = 98$) or CTZ ($n = 96$). A total of 20 patients receiving CTZ and 7 patients receiving placebo achieved remission, defined as a CDAI of ≤ 2.8 at both weeks 20 and 24, and had treatment withdrawn. Only 3/17 CTZ and 2/6 placebo patients maintained remission until week 52, although reinstitution of CTZ induced a renewed improvement. Retreatment with placebo did not occur	The data suggest that CTZ cannot be withdrawn in patients achieving remission
Chatzidionysiou 2012[198]	Sweden	Randomised controlled open-label pilot study evaluating whether or not remission can be sustained after cessation of ADA in patients with a DAS28 of < 2.6 for ≥ 3 months ($n = 31$). Remission was rarely maintained in patients who discontinued ADA. The proportion with sustained remission was significantly lower than among those continuing on ADA	ADA discontinuation may be feasible only in a minority of patients with RA in stable clinical remission

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
Haragai 2012[199]	Japan	A retrospective study to assess the cessation of ADA monotherapy in patients with low disease activity (i.e. a DAS28 of < 2.7). In total, 24 patients continued ADA treatment, with 22 patients ceasing treatment. Of these, 14/22 patients did not restart bDMARDs, with 4/22 maintaining low disease activity for 52 weeks	Some RA patients who have achieved low disease activity can discontinue ADA without increasing disease activity. This should be confirmed in a prospective, randomised study
Kaine 2012[200]	USA	A prospective study ($n = 167$) to assess the temporary interruption of ABT treatment. After a 12-week introduction period, 120 patients were randomised to either placebo or continuation of ABT treatment. Following this, 79 patients on placebo were reintroduced to ABT. A non-significant increase in immunogenicity ($p = 0.119$) was observed following ABT withdrawal. Safety was comparable across treatment schedules	A stop–start schedule of ABT was well tolerated with little impact on safety or efficacy
van der Maas 2012[201]	The Netherlands	An observational cohort study assessing the feasibility of down-titrating or discontinuing IFX. IFX was down-titrated by 3 mg/kg (25% of the original dose) every 8–12 weeks until discontinuation or a flare in patients with a DAS28 of < 3.2 for ≥ 6 months. Flares (i.e. a DAS28 of > 3.2) were treated with the last effective dose of IFX. IFX could be discontinued in 16% of the cohort and down-titrated in 45%. There was no statistically significant difference in patients' quality of life ($p < 0.152$) after down-titrating and mean costs per patient were reduced by €3474	In the majority of patients with a stable DAS28 of < 3.2 and stable IFX treatment, IFX can be down-titrated or discontinued without influencing patients' quality of life, generating a considerable cost saving
Klarenbeek 2011[202]	The Netherlands	To determine the relapse rate and predictors of relapse in patients in sustained clinical remission following the withdrawal of treatment. In total, 115 patients in the BeSt study[203] achieved a DAS28 of < 1.6 for > 6 months and all treatment (including conventional DMARDs) was discontinued. A total of 59/115 patients maintained drug-free remission for a median duration of 23 months; 53/115 restarted treatment as the DAS28 reached > 1.6 (median duration to restarting treatment 23 months); and 3 patients were lost to follow-up. In total, 39/53 people who restarted treatment attained remission within 3–6 months of restarting treatment	Approximately 25% of patients with RA achieved drug-free remission; 46% restarted DMARD monotherapy because of a relapse, the majority of whom again achieved clinical remission without showing radiological progression during the relapse
van den Broek 2011[204]	The Netherlands	Post hoc analyses ($n = 104$) of the BeSt study[203] to identify predictive factors for maintaining low disease activity (i.e. a DAS28 of ≤ 2.4) for 6 months without IFX treatment. Low disease activity was maintained in 52% of patients, with a higher success rate in those initially	Maintaining IFX-free low disease activity was successful in the majority of patients and, of those who did flare, the large majority regained low disease activity following treatment with IFX

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
		treated with IFX. Of those who flared (48%), 84% regained low disease activity with IFX treatment. Predictive factors for requiring IFX treatment were smoking, long IFX treatment duration (≥ 18 months) and shared epitope	
Bejarano 2010[205]	England	Prospective cohort study ($n = 20$) of patients with poor prognosis of RA with < 1 year of symptoms. Patients were randomised to receive IFX or placebo for 1 year; these were then removed and patients were treated with MTX monotherapy alone in accordance with standard clinical care. At 8 years, disease activity data were collected. At follow-up, four patients in the IFX group were in remission compared with none in the placebo group. One patient in the IFX group achieved drug-free remission. Median DAS28 was significantly lower in the IFX group than in the placebo group (2.7 vs. 4.3; $p = 0.02$)	A remission induction regime with IFX for 1 year in early RA can improve long-term clinical outcomes
Saleem 2010[206]	England	Prospective cohort study ($n = 47$) of patients in remission attempting to define markers that are predictive of sustained remission following cessation of bDMARD treatment. Of the 47 patients, 27 had received initial treatment and 20 delayed treatment. Two years after treatment withdrawal, 59% in the initial treatment group and 15% in the delayed treatment group had sustained remission. In the initial treatment group, secondary analyses showed that shorter symptom duration was the only clinical predictor of successful treatment withdrawal. Several immunological parameters were significantly ($p < 0.05$) associated with sustained remission	In patients in remission, short duration of untreated symptoms and immunological parameters are associated with successful withdrawal of bDMARDs
Tanaka 2010[184]	Japan	Prospective cohort study ($n = 114$) analysing the possibility of discontinuing IFX after achieving low disease activity (i.e. a DAS28 of < 3.2). When low disease activity was maintained for > 24 weeks, IFX was withdrawn ($n = 102$). A total of 56 patients (55%) maintained low disease activity at 1 year post IFX treatment; 46 patients were not classed as successful, with IFX being restarted or having a DAS28 of > 3.2 at 1 year	After achieving low disease activity through IFX treatment, the majority of patients remained in this state without IFX treatment
Brocq 2009[207]	France	A prospective cohort study of 304 patients taking a bDMARD for RA. Those who achieved remission (i.e. a DAS28 of < 2.6 for at least 6 months without NSAIDs or prednisolone) ($n = 21$) had their bDMARD removed; this was reinstated following a relapse (i.e. a DAS28 of > 3.2)	A relapse occurred within 12 months in 75% of patients who had ceased bDMARD treatment. Relapsing patients responded well to the resumption of the same bDMARD

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
van der Kooij 2009[208]	The Netherlands	An analysis of patients from the BeSt study[203] with a sustained DAS28 of < 1.6 for at least 6 consecutive months who discontinued treatment at year 3. Of the entire cohort, based on the four treatment strategies in the BeSt study, the percentage of patients in drug-free remission at the end of year 4 ranged from 8% to 18%, with a non-significant ($p = 0.14$) difference across strategies	In patients with recent-onset RA, drug-free remission was achieved in up to 18% given DAS-driven treatment
Nawata 2008[209]	Japan	A prospective cohort of 172 patients with active RA were provided with IFX treatment. After induction and maintenance of clinical remission (i.e. a DAS28 of > 2.6), tapering of corticosteroids and/or NSAIDs was attempted. If clinical remission was maintained for > 24 weeks, discontinuation of IFX was considered. In total, 52 patients met the remission criteria and nine patients discontinued IFX and maintained remission for a mean of 14 months without recurrence. The duration of disease was shorter and the points from Steinbrocker's stage classification were significantly lower in the IFX-discontinued group than in the IFX-continued group	The study indicates that DAS28 could be a good indicator of whether or not to perform a strategic reduction of IFX. The findings imply that early intervention with IFX appears to be advantageous for achieving clinical remission and for discontinuing IFX after clinical remission
van den Bilj 2007[210]	The Netherlands	Analysis from the BeSt study[203] of 120 patients with early RA receiving IFX treatment. Sixty-seven patients had persistent (> 6 months) low disease activity (i.e. a DAS28 of ≤ 2.4) and had IFX tapered and finally withdrawn at a median time of 10 months. Ten patients had IFX withdrawn but experienced a disease flare and resumed IFX treatment after a median of 4 months. This represents a bDMARD-free maintenance success rate of 67/120 (56%) during a 2-year follow-up period	56% of patients with early RA, initially treated with IFX, could discontinue IFX after achieving a DAS28 of < 2.4
Quinn 2005[211]	England	A 12-month double-blind study ($n = 20$) attempting to induce remission with or without IFX treatment in those with symptoms for < 12 months. Of those on IFX ($n = 10$) who responded to treatment ($n = 9$), seven maintained their DAS28, whereas two had increases in DAS28 in the following 52 weeks after withdrawal of treatment. No patient had a functional deterioration as measured by the HAQ after withdrawal of IFX treatment	Functional and quality of life benefits were sustained a year following the withdrawal of IFX treatment

TABLE 13 Results from a review of dose tapering studies within RA (*continued*)

Study	Setting	Summarised details[a]	Summarised authors' conclusions
Buch 2004[212]	England	Prospective 2-year extension of the Anti-TNF Therapy in RA with Concomitant Therapy (ATTRACT) study[213] ($n = 17$). All patients had a flare, with 15 choosing retreatment with IFX, although one stopped treatment because of attempted pregnancy. After retreatment the ACR response was comparable in 12 out of 14 patients and worse in 2 out of 14 patients. No adverse reactions were observed	IFX treatment can be restarted after an interval of several months without any observable problems

BeSt, Behandelstrategieen voor Reumatoide Artritis.
a All studies are assumed to have allowed concurrent cDMARD use unless explicitly stated otherwise.

The potential advantages of using ultrasound for monitoring synovitis

The possibility of a reduced bDMARD burden was not considered by the manufacturers, the Assessment Group or the NICE Appraisal Committee in the recent evaluation of bDMARDs,[40] indicating that tapering of treatment in patients who have responded well may not reflect routine clinical practice. A relevant question is, 'Why is this the case?'. If it is because clinicians are unwilling to risk treatment reductions in patients whose disease may worsen if dosing is amended, then monitoring synovitis in patients such that treatment could be reinstated, or the dosage amended, when disease activity reoccurred could have the potential to be a cost-saving strategy and also one that could be beneficial to patients. In a survey sent to members of the BSR (see *Appendix 1*), 27 out of the 31 (87%) respondents replied that US is used to make decisions regarding RA therapy, with the majority stating that US was used to make decisions whether to start or stop medication or to taper or increase dosages. There was no consensus among the respondents about how often US was used in RA patients, although it appeared that this would likely be dependent on patients' symptoms and the uncertainty around disease activity; a minority of respondents indicated that US would be routinely undertaken on patients at visits to the clinic. Although the data provided by those BSR members responding to the survey are helpful, it is not clear to what extent selection bias could have influenced the results. It cannot be ruled out that the majority of those who chose to respond were advocates of the use of US for monitoring synovitis.

Cost-effectiveness analyses undertaken

The economic evaluation compared the use of US as an adjunct to CE and the use of CE only in determining the most appropriate treatment in RA. This analysis was performed using four scenarios: (1) patients who have been perceived to be clinically stable on a bDMARD and for whom the clinician may consider dose reductions; (2) patients who have been perceived to be clinically stable on a cDMARD and for whom the clinician may consider dose reductions; (3) patients who appear to have disease progression despite bDMARD treatment and for whom the clinician is contemplating amending treatment; and (4) patients who appear to have disease progression despite cDMARD treatment and for whom the clinician is contemplating amending treatment. It should be noted that the groups of patients within these four scenarios do not equate to the entire RA population. If a clinician did not believe that escalation or tapering of treatment was warranted, it was assumed that patients would not have synovitis measured with US.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

All analyses were undertaken using a 1-year time horizon assuming that the decisions faced by clinicians would be recurrent, that is, that patients would continue to be monitored to ensure that the initial treatment decision remained appropriate. Given the short time horizon of the analysis, the results have not been discounted. It was assumed that if a strategy of monitoring synovitis using US was employed then this would be consistently undertaken in all patients who met the criteria of the four scenarios in all future years. In all analyses, incidental benefits and incidental costs, for example those related to detecting a condition that could have been mistaken for a worsening of RA, have been excluded from the simple model.

In the base-case analysis it was assumed that, on average, four US scans would be undertaken per year, at a cost of £226.62 per year; this was adjusted in sensitivity analyses. This frequency of monitoring in those patients for whom a clinician was considering a change of treatment or in patients following a change of treatment was considered prudent to ensure that any increase in synovitis was detected relatively early. Were four US scans per year adopted in clinical practice this could impact on sonographer capacity; if additional resources were required then the cost savings required for monitoring of synovitis with US to be cost neutral would increase.

The analyses presented detail the reductions in average drug use and the number of serious infections that would need to be avoided for US plus CE to have an incremental cost-effectiveness ratio (ICER) of £20,000 per quality-adjusted life-year (QALY) gained and £30,000 per QALY gained[214] relative to CE alone. These two willingness-to-pay thresholds have been selected as they are reported in NICE's *Guide to the Methods of Technology Appraisal 2013*[214] and were current at the time of report writing (2016).

Serious infections were considered as potential adverse events of treatment. It was estimated that a serious infection would be associated with a cost of £1479 and a QALY loss of 0.012.[23] Although serious infections are more typically associated with bDMARDs, the possibility of this adverse event has also been included for patients on cDMARDs. The net monetary benefit of an avoided serious infection was calculated assuming both a £20,000 per QALY gained threshold and a £30,000 per QALY gained threshold. The formula for this is:

$$\text{Assumed cost of a serious infection} + (\text{QALY loss associated with a serious infection} \times \text{cost per QALY gained threshold}). \quad (1)$$

Thus, the net monetary benefit of an avoided serious infection using a £20,000 threshold is £1718 [£1479 + (0.012 × £20,000)] and using a £30,000 threshold is £1838 [£1479 + (0.012 × £30,000)].

The threshold levels for reductions in average drug use and the number of serious infections that need to be avoided are not intended to imply that monitoring synovitis with US is or is not cost-effective, rather that these are the levels at which the use of US becomes cost neutral. These threshold values may or may not be achieved in reality and it is conceivable that there would not be a saving if the use of US led to overtreatment.

It is acknowledged that the decision to amend treatment is likely to be multifactorial. However, there is not sufficient certainty in these factors, and potential heterogeneity of their implementation to allow a meaningful analysis to be conducted. We have therefore chosen to subsume all factors into threshold metrics (either in reduction of drug use or in reduction in serious infection) that would need to be achieved through the extra information supplied by the use of US as an adjunct to CE.

Analyses undertaken when dose tapering is being considered

For the modelling of strategies in which a clinician is contemplating a dose reduction, a simplistic approach has been taken using threshold analyses. This is in contrast to the protocol,[215] in which it was anticipated that the model constructed for the recent NICE appraisal of bDMARDs and documented in Stevenson *et al.*[23] would be used. Given the uncertainty in the evidence base, it was deemed more useful to provide indicative results that focus on the key parameters related to US use, rather than to provide results of potentially spurious accuracy from a model with a much larger number of parameters, which could obscure the impact on the decision problem that we have been tasked to address.

Analyses undertaken when a change in treatment is being considered

The potential consequences associated with changing treatment are more complicated than those associated with tapering treatments as patients may have a decrease in HAQ score if they respond to the new intervention. More aggressive treatment could also potentially stop disease progression. However, as with the analyses relating to dose tapering, a simplistic approach has been undertaken to provide an indicative level of the proportion of patients not escalating treatment, because of the added information provided by a US scan, that would be result in an ICER of £20,000 or £30,000 per QALY gained. The costs of DMARDs and the impacts of a serious infection have been assumed to be the same as in the tapering analyses. Incidental benefits, for example detecting a condition that could have been mistaken for a worsening of RA, have been excluded from the model.

Costs assumed within the model

As detailed in *Chapter 3* (see *Anticipated costs associated with the intervention*), we calculated that the use of US to monitor synovitis would be associated with a cost of £56.66. We assumed that all US appointments were outpatient appointments, that contrast was not used and that two-thirds of appointments took < 20 minutes and the remainder took ≥ 20 minutes.

For simplicity, it was assumed that clinical assessment was associated with no cost; as clinical assessment is included in both treatment strategies, and there is unlikely to be a mortality difference between strategies, this was believed not to influence the results.

The assumed cost of a bDMARD was sourced from a recent *Health Technology Assessment* report.[23] Several bDMARDs are subject to a commercial-in-confidence patient access scheme; however, for the interventions for which an estimated cost per year can be reported, the prices were commonly around £9200 per year. This price was assumed in our analyses. It should be noted, however, that biosimilar drugs have entered the market for IFX, ADA, RTX and ETN and thus, potentially, drug acquisition costs may decrease in future years. An additional £1608 per year was assumed for monitoring and administration for all regimens.

The assumed cost of cDMARD treatment was dependent on the actual regimen used. The most intensive course of treatment that would be used was assumed to replicate that in the study by Stevenson *et al.*:[23] oral MTX (20 mg weekly), HCQ (6.5 mg/kg daily), SSZ (3 g daily) and oral prednisolone (7.5 mg daily). This had an estimated cost of £1826 per year when monitoring and administration costs were included and £218 per year when these costs were excluded.

In a sensitivity analysis the impact of MTX being given subcutaneously rather than orally was explored, as this is the formulation recommended for people with severe active RA, with oral MTX recommended for people with moderate to severe RA.[216] The annual cost of two 10-mg tablets of MTX per week is £39.28 whereas the annual cost of a weekly 20-mg solution for injection is £837.99.[216] The costs of intensive cDMARDs and MTX were both increased by £798.71 (£837.99 – £39.28) in the sensitivity analysis.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The costs of serious adverse advents are included; these were estimated to be £1479.[23]

Utilities assumed within the model

The only utility implication considered within the model was that associated with serious adverse events. These were assumed to be associated with a QALY loss of 0.012.[23]

Summarised model inputs

Table 14 provides the model parameters used in the base-case analysis.

An illustrative example of how thresholds were calculated

An example of the calculation of a threshold is provided, assuming a cost per QALY gained threshold of £20,000 and a reduction in the costs of bDMARDs of 1%. A saving of £92 (£9200 × 1%) would be made on bDMARD acquisition costs, which would reduce the net costs of a US strategy to £134.62 (£226.62 – £92). Given a net monetary benefit of £1718 per serious infection avoided, 0.078 (£134.62/£1718) serious infections would need to be avoided to achieve a cost per QALY gained of £20,000.

TABLE 14 Model parameters used in the base-case analysis

Parameter	Value	Reference
Cost of a US scan	£56.66	Assumption based on *NHS Reference Costs 2014 to 2015*[73]
Number of US scans per year per patient	4	Assumption
Cost of US scans per year per patient	£226.62	Calculated
Annual cost of bDMARD treatment[a]	£9200	Stevenson *et al.*[23]
Annual cost of intensive cDMARD treatment[a,b]	£218.17	BNF[216]
Annual cost of oral MTX treatment[a]	£39.28	BNF[216]
Annual cost of subcutaneous MTX treatment[a]	£837.99	BNF[216]
Cost of a serious infection	£1479	Stevenson *et al.*[23]
QALY loss associated with a serious infection	0.012	Stevenson *et al.*[23]
Net monetary benefit of a serious infection avoided at a threshold of £20,000 per QALY gained	£1718	Calculated
Net monetary benefit of a serious infection avoided at a threshold of £20,000 per QALY gained	£1838	Calculated

BNF, *British National Formulary*.
a Excluding monitoring and administration costs, which were estimated to be £1608 per annum.
b Assumed to consist of MTX (20 mg weekly), HCQ (6.5mg/kg daily), SSZ (3 g daily) and oral prednisolone (7.5 mg daily).

Results

The cost-effectiveness of ultrasound monitoring in patients who have been stable on biological disease-modifying anti-rheumatic drugs and for whom the clinician is contemplating reducing the dose of biological disease-modifying anti-rheumatic drug

The average reduction in bDMARD use that would be required for a strategy of US monitoring to be cost neutral was calculated. This was 2.46% (£226.62/£9200). The levels of drug reduction and serious infections avoided at which US is at the threshold of cost-effectiveness are shown in *Figure 2*.

The cost-effectiveness of ultrasound monitoring in patients who have been stable on conventional disease-modifying anti-rheumatic drugs and for whom the clinician is contemplating reducing the dose of conventional disease-modifying anti-rheumatic drug

The average reduction in cDMARD use that would be required for a strategy of US monitoring to be cost neutral was calculated. It was not possible to recoup the assumed costs of US (£227) with reduced use of either intensive cDMARDs (£218 per annum) or MTX (£39 per annum). However, the levels of drug reduction and serious infections avoided at which US is at the threshold of cost-effectiveness are shown in *Figure 3*.

The cost-effectiveness of ultrasound monitoring in patients who appear to have disease progression despite biological disease-modifying anti-rheumatic drug treatment and for whom the clinician is contemplating amending treatment

This analysis was not formally conducted as there was insufficient evidence to provide any robust assessment. The likely impact of a change in treatment for a patient on bDMARDs will vary depending on the current treatment and on whether or not the disease was progressing.

At the time of writing, NICE guidance[41] recommended that patients progressing from first-line bDMARDs should receive RTX in combination with MTX as it is cheaper than other bDMARDs and has a similar efficacy.[23] As such, the immediate impact of a patient moving to RTX would be a cost saving, regardless of whether or not the disease was progressing and regardless of any US result. However, a potential treatment option has been exhausted, which may have longer-term impacts if the disease had not progressed at the time of the switch to RTX, but began progressing on subsequent treatments. It is unclear to what extent the decision to move to RTX would be influenced by the use of US to monitor synovitis.

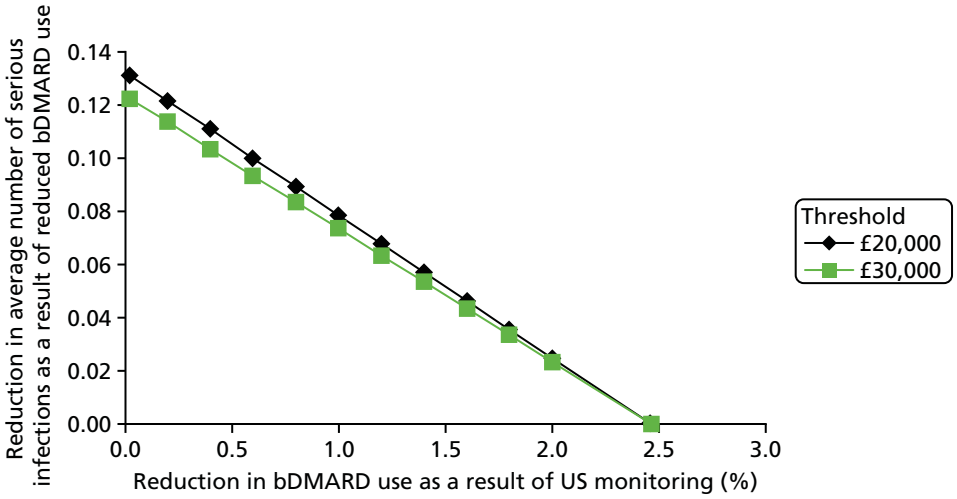


FIGURE 2 The average levels of bDMARD drug reduction and serious infections avoided that result in a cost per QALY gained of £20,000 and £30,000 for the use of US in monitoring synovitis.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

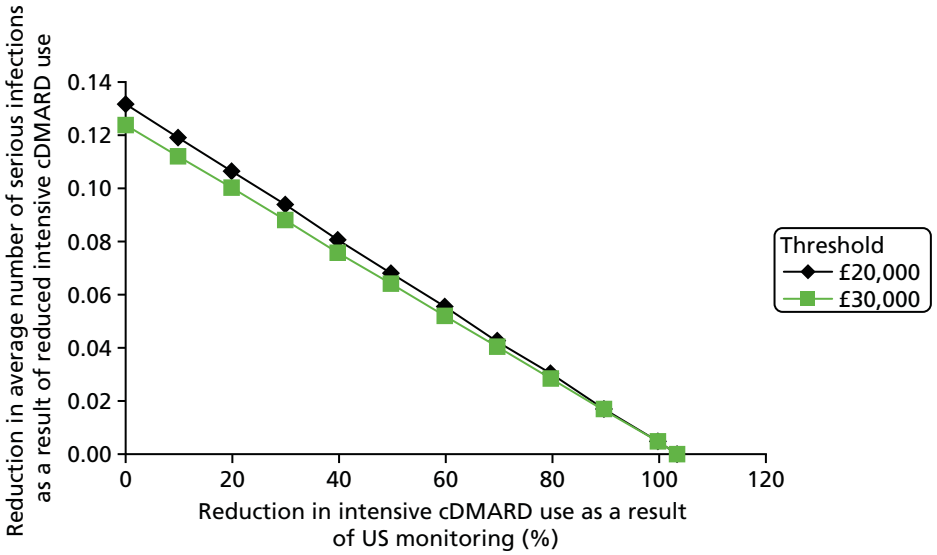


FIGURE 3 The average levels of intensive cDMARD drug reduction and serious infections avoided that result in a cost per QALY gained of £20,000 and £30,000 for the use of US in monitoring synovitis.

In contrast, if a patient was on full-dose TCZ following RTX treatment, at the point that the clinician was contemplating amending treatment, current NICE guidance would allow no further bDMARDs. It is unclear to what extent the knowledge of underlying synovitis would influence a clinician's decision to amend treatment. The threshold proportion of patients who could be moved onto RTX from their initial bDMARD, at which point the savings in bDMARD acquisition cost would offset the costs of US monitoring, has been calculated. This threshold value, assuming a cost per year of £4900 for RTX, would be £226.62/(£9200 – £4900), which is 5.3%. The components of this formula are the annual costs of US monitoring and the annual costs associated with bDMARDs and RTX.

If a patient was on a reduced-dose bDMARD because of previous tapering then the use of US to decide whether or not to increase the dose is easier to model, although the answer would be dependent on the assumed increase in dose. It has earlier been shown that the assumed costs of US per year equate to 2.46% of the annual costs of bDMARDs. The proportion of patients who would not need an increase in dose because of the information provided by US for a strategy of monitoring with US to be cost neutral can be calculated as £226.62/[assumed increase in bDMARD dose (as a percentage of the full dose) × £9200].

Thus, if a dose increase of 33% of the full dose was contemplated, the threshold proportion of patients who do not need to be treated would be £226.62/(33% × £9200), which is equal to 7.4%.

The cost-effectiveness of ultrasound monitoring in patients who appear to have disease progression despite conventional disease-modifying anti-rheumatic drug treatment and for whom the clinician is contemplating amending treatment

This analysis has again been divided into those patients on intensive cDMARDs and those on MTX alone. It has been assumed that those on MTX alone would progress to an intensive cDMARD strategy, whereas those on intensive cDMARDs would move to a bDMARD. The thresholds for patients on intensive cDMARDs are shown in *Figure 4* and the thresholds for those on MTX alone are shown in *Figure 5*. Assuming no changes in serious infections the threshold value was 2.52% for those not progressing to bDMARDs; the threshold value could not be attained for those not progressing to intensive cDMARDs.

When a patient is on a reduced dose of intensive cDMARDs or MTX because of previous tapering, the threshold for the proportion of patients at which monitoring with US becomes cost neutral can be calculated in a similar manner to that illustrated for patients on bDMARDs who have received tapered medication and the clinician is contemplating increasing the dose.

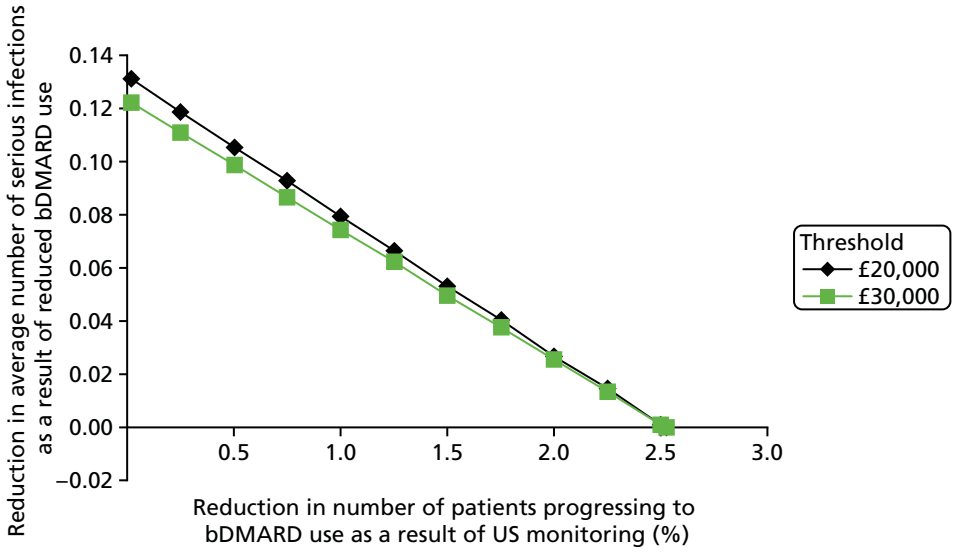


FIGURE 4 The average levels of reduction in patients moving to bDMARDs and serious infections avoided that result in a cost per QALY gained of £20,000 and £30,000 for the use of US in monitoring synovitis.

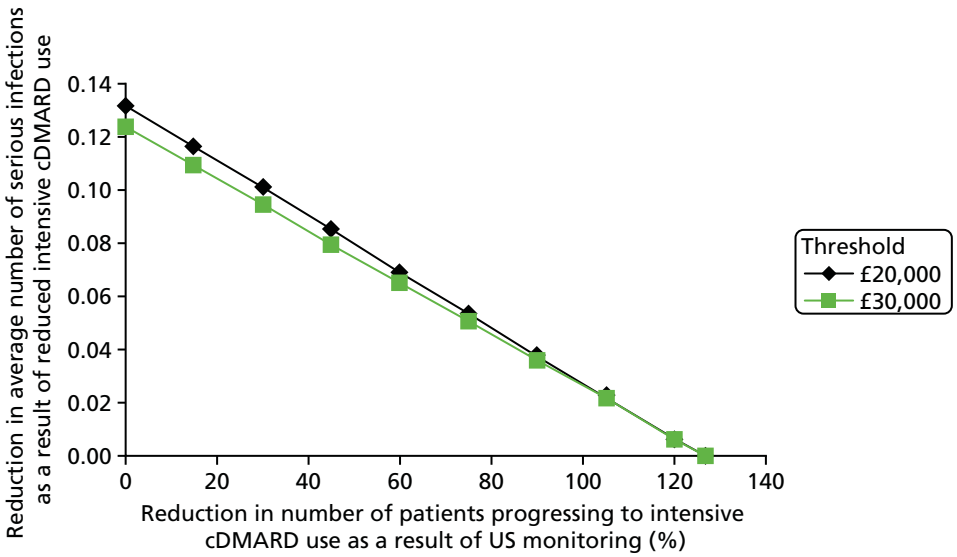


FIGURE 5 The average levels of reduction in patients moving to intensive cDMARDs and serious infections avoided that result in a cost per QALY gained of £20,000 and £30,000 for the use of US in monitoring synovitis.

Sensitivity analysis

Sensitivity analysis was undertaken assuming that the number of US scans required per patient to monitor synovitis ranged from one per year to 12 per year. The cost of a scan was assumed constant at £56.66.

Figure 6 provides data relating to the analyses in which the treatment dose was being tapered. It was calculated that, even if a scan was provided on a monthly basis, this approach would be cost neutral if the average decrease in bDMARD costs was 7.39%. This value was 15.99% for intensive cDMARDs. In contrast, for patients on intensive cDMARDs, even when only one scan was performed per year a reduction in drug costs of 26% would be needed to make US cost neutral.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

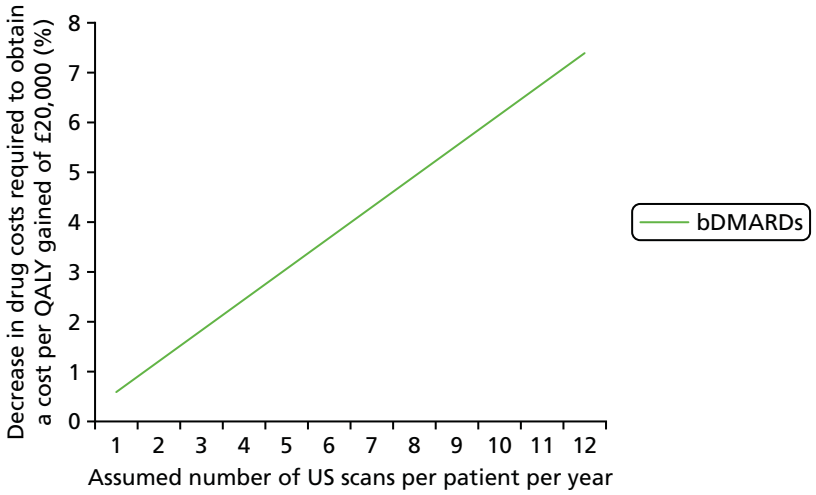


FIGURE 6 Sensitivity analysis relating reduction in drug-related costs to the assumed number of US scans undertaken per year.

Further sensitivity analysis was undertaken looking at the percentage reduction in the cost of bDMARDs, assuming lower prices than the £9200 assumed in the base case. The motivation for this sensitivity analysis was the fact that biosimilars have entered the market. This sensitivity analysis is shown in *Figure 7*.

It is seen that, as the price of bDMARDs approaches 50% of the price assumed in the base case, a 5% reduction in bDMARD use would still be sufficient to make monitoring with US cost saving.

Additional sensitivity analysis was conducted to assess the impact of the increased costs associated with use of subcutaneous MTX, which was assumed to cost £798.71 per annum more than oral MTX.

The results for the tapering analyses when assuming costs associated with subcutaneous MTX are provided in *Figures 8–11*. The reduction in drug costs at which the use of US became cost neutral were 2.27% for patients on bDMARDs, 22.9% for those on intensive cDMARDs and 27.04% for patients on subcutaneous MTX alone.

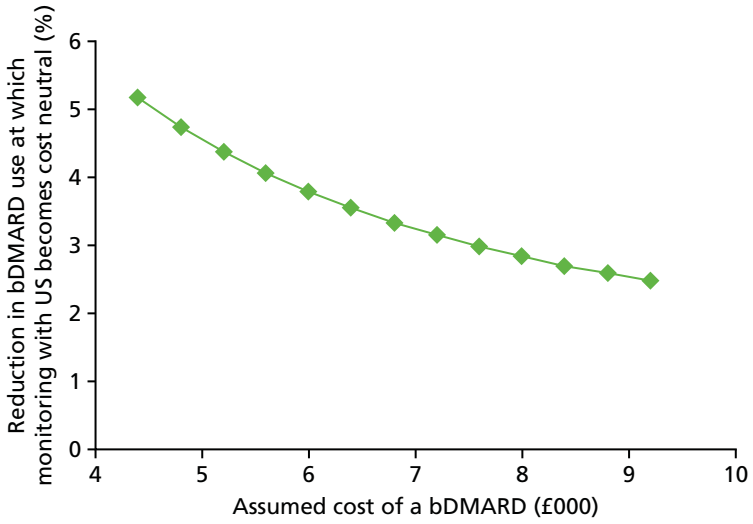


FIGURE 7 Sensitivity analysis relating percentage reduction in bDMARD acquisition costs to the assumed price of bDMARDs.

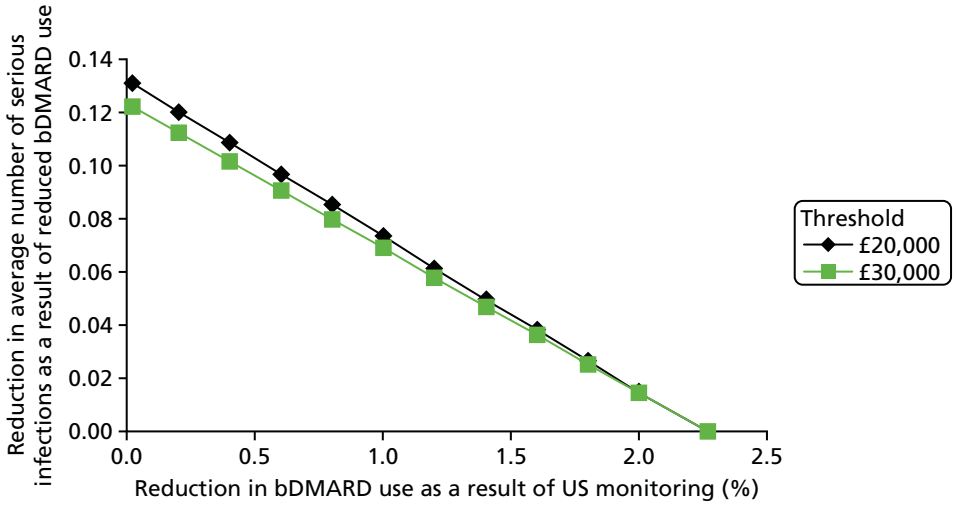


FIGURE 8 The average levels of bDMARD drug reduction and serious infections avoided that result in a cost per QALY gained of £20,000 and £30,000 for the use of US in monitoring synovitis assuming the use of subcutaneous MTX.

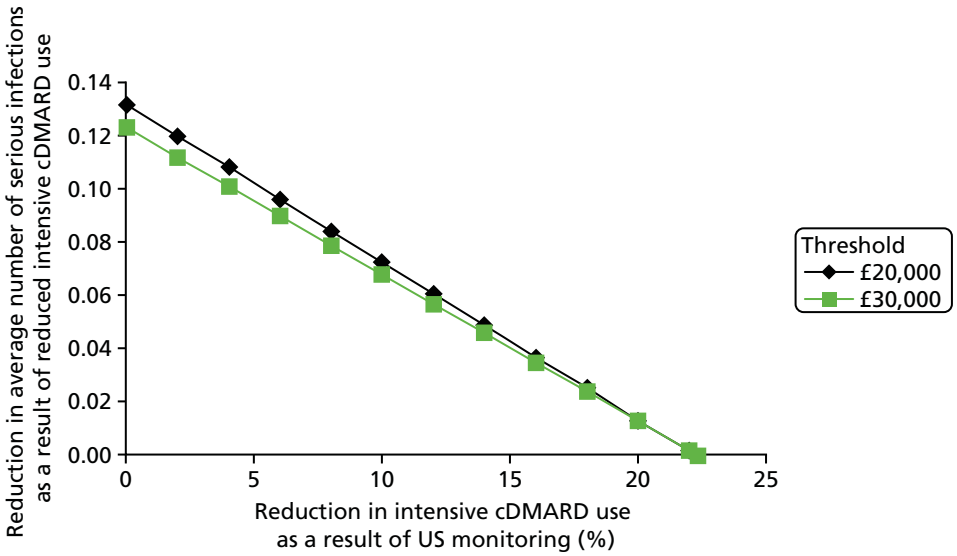


FIGURE 9 The average levels of intensive cDMARD drug reduction and serious infections avoided that result in a cost per QALY gained of £20,000 and £30,000 for the use of US in monitoring synovitis assuming the use of subcutaneous MTX.

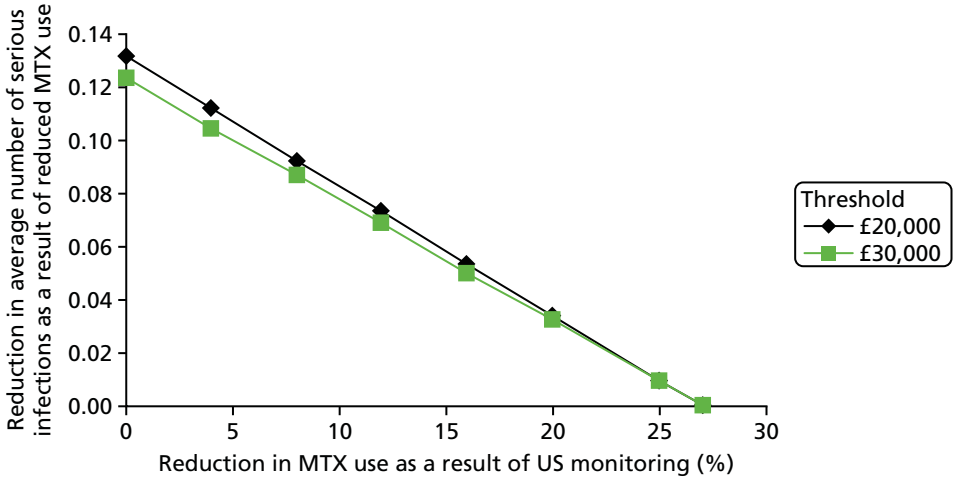


FIGURE 10 The average levels of MTX drug reduction and serious infections avoided that result in a cost per QALY gained of £20,000 and £30,000 for the use of US in monitoring synovitis assuming the use of subcutaneous MTX.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

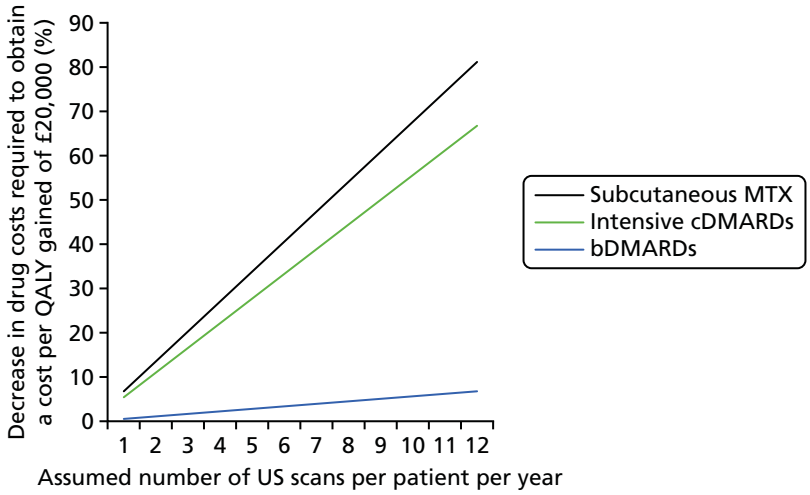


FIGURE 11 Sensitivity analyses relating reduction in drug-related costs to the assumed number of US scans undertaken per year assuming the use of subcutaneous MTX.

Interpretation of the results

For patients for whom a clinician is contemplating a reduction in bDMARD use because of stable disease, our analyses indicate that only a small average reduction in bDMARD use (2.46%) would be sufficient to offset the costs associated with monitoring patients with US. This value is likely to be higher than in reality should the drug burden be reduced for multiple years and the frequency of monitoring be reduced. Data collated from a review of the evidence for bDMARD tapering (see *Table 13*) indicate that dose reduction can occur in a sizeable proportion of patients. As such, a 2.46% reduction in bDMARD dose, which is sufficient to offset the costs associated with monitoring patients with US, could be seen to be plausible if the use of US made a clinician more confident to taper bDMARDs. An analysis was undertaken exploring how these results would change if the average price of bDMARDs falls following the introduction of biosimilars. This indicated that, if the cost was to fall by 50% of that assumed in the base case, the reduction in bDMARD costs required for the use of adjunct US to be cost neutral would remain below 5%.

Reducing intensive cDMARD drug costs or MTX costs alone would not be sufficient to offset the costs of US.

For both the bDMARD and cDMARD tapering scenarios there is the potential for fewer serious adverse events if drug intake is reduced. This has been explicitly shown in *Figures 2–4*. The calculations suggest that 0.13 serious infections saved per person (or approximately one per eight patients) would be sufficient to make US for monitoring synovitis cost-effective. However, this threshold may be academic only: it is not known whether or not such reductions in adverse events would be attainable only with dose reductions that would already have made monitoring with US cost saving.

For patients who appear to undergo disease progression despite bDMARD treatment and for whom the clinician is contemplating amending treatment, the decision problem is more complex. No analyses were undertaken for patients changing treatment because of the uncertain impacts of downstream treatment. However, an example calculation of the threshold value for the number of patients in whom increased treatment was not necessary was provided for patients increasing their dosage subsequent to previous tapering. The threshold value is dependent on the proposed increased dose of the bDMARD and it is uncertain whether or not US is likely to be cost-effective in this group.

The use of US in monitoring synovitis in patients on cDMARDs was modelled as there were clear next-treatment options dependent on the current regimen, namely bDMARDs for those on intensive cDMARDs

or intensive cDMARDs for those on MTX alone. The results in this group were not dissimilar from those when tapering was being considered, with a relatively small (approximately 2.5%) reduction in the number of patients progressing to bDMARDs being required to make the use of US cost-effective.

Ultimately, the cost-effectiveness of US for monitoring synovitis will be dependent on how clinicians change their decisions about care given the additional knowledge provided by US. If the results do not feed into treatment strategies, then the use of US would not be cost-effective. However, evidence presented in *Table 12* suggests that the use of US to monitor synovitis was fairly prevalent, with a range of 23–88% for studies based in the UK.[89,90,153] Furthermore, it was reported that clinician confidence was significantly improved ($p < 0.0005$) by the use of US.[152] As such, it is anticipated that the use of US in monitoring synovitis is plausibly cost-effective, particularly when activity identified on US has been shown to be predictive of flares.[142,155,156,163,164]

However, the authors note that there remains considerable uncertainty around all decisions because of the lack of robust evidence and that any conclusion should be interpreted with caution.

Analysis was undertaken to assess the impact of the increased acquisition price if subcutaneous MTX were used rather than oral MTX. This reduced the thresholds required for the use of adjunct US to be cost neutral, but there remained considerable uncertainty in the result.

Discussion

The modelling undertaken was purposefully simplistic so that the key interactions between monitoring synovitis with US and decisions influencing treatment could be examined explicitly. The results presented appear to indicate that monitoring synovitis with US could be cost-effective given the evidence on the potential tapering of DMARDs presented in *Table 13*.

It has been assumed that monitoring with US would allow a clinician to better identify disease worsening when drugs have been tapered. If that is not the case, and irreversible damage has been caused to patients, then the thresholds presented will be favourable to US monitoring, with patients potentially costing more in direct medical costs and having a reduced health-related quality of life. Although evidence on long-term patient outcomes is limited, the majority of comparative studies have reported that there were no differences in outcomes between tapered groups and non-tapered groups and comparative and observational data have shown that patients who flared regained a favourable disease activity state when treatment was reintroduced. Evidence for successful ADA or CTZ tapering is arguably less strong than evidence for successful IFX tapering, although longer-term data are needed to form any conclusions.

If it is possible to reduce the frequency of US monitoring post DMARD tapering, or having established that treatment escalation is not required, then the threshold values presented in this report will be unfavourable to US monitoring.

Data from randomised clinical trials assessing the monitoring of synovitis with US as an adjunct to CE compared with CE alone are needed. The TaSER study[78] (see *Table 11*) was one such RCT and the full results of this study were included in the review described in *Chapter 3*. Incorporating long-term radiographic data into such a study would allow an insight into whether or not irreversible damage is caused by tapered treatment. Wakefield *et al.*[24] state that 'there is now a compelling argument to suggest that the addition of ultrasound assessment . . . is likely to improve the prediction on clinical outcomes'. At the time of writing, ARCTIC trial interim data had been presented in an abstract[79] (November 2015) and these data were included in *Chapter 3*. At the time of the last search, the Targeted Ultrasound in Rheumatoid Arthritis (TURA) study[217] was ongoing and had a reported 32% global recruitment rate. As an identified ongoing study, the authors checked (in February 2016) for any publications from the TURA study; however, it was still ongoing. There is no information available on when these studies will report results. The objective of the ARCTIC study

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

was to assess whether or not the information provided by ultrasonography assessment produces better outcomes than when treatment decisions are made solely on clinical and laboratory assessments. The results from these trials are expected to reduce the decision uncertainty regarding whether or not monitoring of synovitis in RA patients with US is clinically effective and cost-effective. Furthermore, two ongoing studies (NCT02064400 and NCT01602302) were identified from a search of ClinicalTrials.gov, although neither is a RCT.

Chapter 5 Assessment of factors relevant to the NHS and other parties

Evidence identified for this report appears to indicate that monitoring synovitis with US could potentially be cost-effective given the possibility of tapering of DMARDs or inappropriate escalation of treatment. There are a number of potential implications for NHS resources. If monitoring of synovitis becomes more widespread (and there are a number of unknowns, for example US use may vary according to the apparent clinical level of disease activity), depending on the number of US scans undertaken per year, there may be a need for more sonographers. Such increased resources would also depend on drivers outside the scope of this report. US has clinical advantages beyond the monitoring of inflammatory arthritis, both for diagnosis (e.g. for a proportion of patients with suspected inflammatory arthritis, enabling earlier intervention) and for guiding the accurate placement of needles for therapeutic injection. All of this has implications for US access in both rheumatology and radiology departments. This also raises issues regarding timely and equal access to US.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 6 Discussion

Principal findings

Studies included in the systematic review varied in interventions, comparators and outcomes, making it difficult to draw firm conclusions from the available evidence. Fifteen studies provided prognostic data.[55,69,97,98,113,134,135,137–144] Although the study designs were of high quality (prospective cohort), the data reported were correlations and sample sizes ranged from 10 to 453 (in total, data were available from 1523 patients but study heterogeneity precluded meta-analysis). PDUS was significantly associated with radiographic progression at follow-up in all studies in which it was investigated.[55,69,97,113,135,138–140,144] The majority of studies investigating radiographic progression reported that GSUS was a significant predictor; however, this was not the case in all studies of GSUS. Some studies found a significant association between DAS and later radiographic progression. PDUS was significantly associated in the majority of studies with follow-up DAS28,[55,98,137] ACR/EULAR remission[145] and flare.[142]

Nine studies reported data regarding the use of US for monitoring treatment response or strategies.[79,93,95,100,101,139,154–156]

Two were RCTs[79,154] and the others were prospective cohort studies.[93,95,100,101,139,155,156] Most studies relied on one-off US measurements at baseline; however, the two RCTs employed US serially as part of the treatment strategies. Sample sizes ranged from 24 to 130, with data available from 627 patients in total (although study heterogeneity precluded meta-analysis). These studies found that US predicted treatment continuation and response to treatment. When treatment was discontinued or tapered, PDUS significantly predicted relapse or disease flare and was a better predictor than DAS28. In the RCTs, the addition of PDUS to a DAS28-based treatment strategy did not lead to significant between-group differences in change from baseline in DAS44 and RAMRIS erosion score[154] or a significant between-group difference in the primary end point, which consisted of a DAS of < 1.6 and no swollen joints at 16, 20 and 24 months and no progression in vdHSS between 16 and 24 months.[162] However, in the TaSER study, the US plus DAS group had significantly more patients attaining DAS44 remission. Six small studies (sample size 17–109 patients; 321 patients in total) reported observational data on the impact of US on treatment decisions.[89–91,94,152,153] Adding US to CE led to modified treatment decisions and significantly improved clinician confidence in these decisions.

A simple model was constructed. Monitoring with US was not assumed for all patients but rather for those for whom the clinician was contemplating a change (escalation or tapering) of treatment. Screening of all patients could have wider implications for cost-effectiveness and was not considered here. The modelling undertaken indicated that only small proportions of patients tapering treatment, or not escalating treatment, were necessary for the savings associated with reduced treatment to offset the costs associated with monitoring synovitis with US. A systematic review of the literature on tapering of DMARDs indicated that some patients achieving low disease activity could have treatment tapered with no, or little, resultant harm, with the majority of those who flared regaining their previous condition on retreatment. As such, there is clearly potential for the monitoring of synovitis with US to be cost-effective.

Strengths, limitations and uncertainties

Strengths

Relevant literature on prognosis, treatment and cost-effectiveness was reviewed systematically according to a prespecified protocol. The research team was independent and experienced in the methodology and clinical advisors were experienced in the field. A patient advisor provided information on patient experience and contributed to the plain English summary.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The review findings agreed with the findings of previous reviews with regard to diagnosis and prognosis. We also compared US with CE in relation to treatment strategies/decisions, which had not previously been systematically reviewed.

Results from this study are generalisable to UK practice. The studies included in the review used international definitions of RA; most used standardised semiquantitative scoring systems for US measures of synovitis; and clinical measures were also used internationally, for example the DAS28. Treatments assessed included cDMARDs and bDMARDs, as would be relevant to UK practice.

To our knowledge, this is the first study to model the cost-effectiveness of the use of US to monitor synovitis. In addition, we also carried out a systematic review of the literature on the tapering of DMARDs.

Limitations/uncertainties

There is no conclusive gold standard/reference standard for assessing synovitis and so diagnostic accuracy comparisons may be flawed. The heterogeneity of the trials precluded meta-analysis; therefore, no summary estimates of effect were available, which is a limitation of the review. Few studies were identified that compared US with CE with regard to their effect on treatment. Few data were identified regarding the additional influence of US in current practice. The review excluded studies not published in English. The inclusion of conference proceedings may have resulted in the effectiveness of the interventions being underestimated.[218]

There is a lack of long-term follow-up of studies of tapering medication. The survey conducted had few responses and was subject to selection bias.

The mathematical model was simplistic, although this was by design to enable us to focus on aspects clearly related to the impact of monitoring synovitis with US in patients with RA. There remains considerable uncertainty around the cost-effectiveness of the use of US for monitoring synovitis.

The systematic review of the literature on DMARD tapering had relatively poor sensitivity, with the majority of papers being identified from reference searching or being known to our clinical advisors. As such, relevant papers may not have been identified, although it is unlikely that these would change the broad conclusions of the review.

Chapter 7 Conclusions

Little evidence matched the decision problem. However, this systematic review found correlational evidence that US-detected synovitis was significantly associated with later radiographic progression. PDUS-detected synovitis also significantly predicted DAS28, ACR/EULAR remission and flare. Studies suggested that US was superior to CE alone in predicting response to treatment tapering or discontinuation. Studies varied with regard to interventions, comparators and outcomes, precluding meta-analysis and meaning that there is uncertainty in the available evidence. With regard to treatment strategies, studies were not all consistent with regard to the statistical significance of the benefit of US. Only two RCTs were identified and the addition of US to DAS-based strategies did not significantly influence the primary end points, although one of these RCTs significantly favoured the addition of US for the outcome of attaining DAS remission. It is uncertain whether the lack of significant influence on the primary end point is due to a genuine lack of improvement caused by the US strategy or other factors such as joint counts and types of US used within studies. Small studies suggested that US was used in treatment decisions and could increase physician confidence in those decisions. A small survey of UK rheumatology units suggested that US is already being used in some units for modifying treatment decisions in RA. Most studies relied on one-off US measurements at baseline; however, the two RCTs employed US serially as part of treatment strategies.

The purposefully simplistic modelling undertaken shows that there is potential for the monitoring of synovitis with US to be cost-effective, although there remains considerable uncertainty around this conclusion.

Implications for service provision

Limited evidence identified for this report appears to indicate that monitoring synovitis with US in those patients for whom clinicians are contemplating a change in treatment has the potential to be cost-effective, given the possibility of tapering of DMARDs or inappropriate escalation of treatment. There are a number of potential implications for NHS resources. If monitoring of synovitis becomes more widespread (and there are a number of unknowns, e.g. US use may vary depending on the apparent clinical level of disease activity), there may be a need for more sonographers, depending on the number of US scans undertaken per year. Such increased resources would also depend on drivers outside the scope of this report. US has clinical advantages beyond the monitoring of inflammatory arthritis, both in diagnosis (e.g. for a proportion of patients with suspected inflammatory arthritis, enabling earlier intervention) and in guiding the accurate placement of needles for therapeutic injection. All of this has implications for US access in both rheumatology and radiology departments. This also raises issues regarding timely and equal access to US.

Suggested research priorities

An important future research recommendation is to evaluate the role of US using methodologically robust studies that prospectively evaluate test accuracy and evaluate the role of US. For any future study thought should be given to validity and efficient design, as discussed by Bossuyt *et al.*,[219] who emphasise the benefits of retaining solely those patients who have discordant results (in this case, between US and CE) and randomising these patients to amended treatment or usual treatment. A feature of the decision problem in this study, however, is the temporal aspect: it could be that clinicians were inclined to taper treatment but would wait to be further convinced and would have amended treatment earlier if their view was supported by evidence from US.

The heterogeneity of trials precluded meta-analysis; therefore, no summary estimates of effect were available, which is a limitation of the review. Heterogeneity could be limited in future studies by employing protocols that use comparable US joint counts (data would have to be presented for all joints scanned so

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

that comparisons could be made with the currently used joint sets, which vary substantially in the included joint count); comparable scoring systems for GSUS, PDUS and GSUS and PDUS combined, including similar severity scales; comparable US definitions of what constitutes 'positive' involvement at both the joint and the patient level (important for trials in which positive/negative findings dictate therapeutic decision-making); and the same clinical outcomes (which has been less of a problem in more recent studies that have reported a range of clinical outcomes).

Further important research questions include:

- What are the long-term effects on function and joint damage in patients who have had treatment tapered as a result of imaging findings?
- Which joints should be assessed with US?
- Can a pre-selected set of joints be used for imaging or does imaging need to be guided by symptoms?
- How often should US assessment occur and could it be restricted to situations in which a treatment decision is to be made, which should be assessed in serial testing studies?
- Would the cost-effectiveness of US change if all patients were screened rather than only those patients for whom a change in management was being considered?

Studies including an assessment of costs and health-related quality of life would be useful to inform future health policy decisions.

Given the range of questions that still need answering, it is unlikely that one study could provide all of the answers. Different frequencies of monitoring could be compared within a randomised trial; however, feasibility or trial practice would limit how many different frequencies of monitoring could be assessed within one trial. Similarly, there are many variations of joint counts. The more joints assessed, the more time is needed for assessment. Different joint counts need to be assessed to determine the most efficient way to use US; enough joints need to be measured to warrant treatment decisions being influenced, while ensuring that the time required by sonographers and patients remains manageable. It is uncertain whether or not particular joint counts could serve all patients or whether or not the most affected joints of individual patients would need to be taken into account. Similarly, there are different systems for assessing the severity of US pathology. The OMERACT initiative has developed and validated standardised scoring systems; however, trials would need to assess, for example, whether GSUS or PDUS alone could be used to inform treatment strategies or whether or not a combination of GSUS and PDUS would be more effective.

Acknowledgements

We would like to thank Mike Holmes [School of Health and Related Research (ScHARR), University of Sheffield] for help with sifting the cost-effectiveness searches, Eva Kaltenthaler and Paul Tappenden (ScHARR) for providing comments on the draft report and Gill Rooney (Programme Manager, ScHARR) for providing administrative support and preparing and formatting the report.

We would like to thank the patient advisor who contributed to the report and the National Rheumatoid Arthritis Society (NRAS) who provided names of patients who were willing to be involved in the project. We would like to thank the BSR for providing input into the survey to investigate how US was being used in practice and for publicising it to UK rheumatology units.

This report was commissioned by the National Institute for Health Research (NIHR) HTA programme. The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA programme.

Contributions of authors

Emma Simpson (Research Fellow) led the project and conducted the systematic review.

Emma Hock (Research Fellow) conducted the systematic review.

Matt Stevenson (Professor of Health Technology Assessment) conducted the cost-effectiveness review and the economic modelling.

Ruth Wong (Information Specialist) conducted the literature searching.

Naila Dracup (Information Specialist) conducted the literature searching.

Allan Wailoo (Professor of Health Economics) commented on the modelling.

Philip Conaghan (Professor of Musculoskeletal Medicine) provided clinical advice throughout the project.

Cristina Estrach (Consultant Rheumatologist) provided clinical advice throughout the project.

Christopher Edwards (Professor of Clinical Rheumatology) provided clinical advice throughout the project.

Richard Wakefield (Senior Lecturer/ Honorary Consultant in Rheumatology) provided clinical advice throughout the project.

Data sharing statement

The data from the systematic review can be obtained from the corresponding author on request.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

References

1. Scott DL, Steer S. The course of established rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2007;**21**:943–67. https://doi.org/10.1016/j.berh.2007.05.006
2. Pincus T, Fuchs HA, Callahan LF, Nance EP Jr. Kaye JJ. Early radiographic joint space narrowing and erosion and later malalignment in rheumatoid arthritis: a longitudinal analysis. *J Rheumatol* 1998;**25**:636–40.
3. Drossaers-Bakker KW, Kroon HM, Zwinderman AH, Breedveld FC, Hazes JM. Radiographic damage of large joints in long-term rheumatoid arthritis and its relation to function. *Rheumatology* 2000;**39**:998–1003. https://doi.org/10.1093/rheumatology/39.9.998
4. Allaire S, Wolfe F, Niu J, LaValley MP, Zhang B, Reisine S. Current risk factors for work disability associated with rheumatoid arthritis: recent data from a US national cohort. *Arthritis Rheum* 2009;**61**:321–8. https://doi.org/10.1002/art.24281
5. Naz SM, Symmons DP. Mortality in established rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2007;**21**:871–83. https://doi.org/10.1016/j.berh.2007.05.003
6. Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, Fautrel B. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. *Joint Bone Spine* 2013;**80**:29–33. https://doi.org/10.1016/j.jbspin.2012.02.005
7. Meune C, Touze E, Trinquart L, Allanore Y. Trends in cardiovascular mortality in patients with rheumatoid arthritis over 50 years: a systematic review and meta-analysis of cohort studies. *Rheumatology* 2009;**48**:1309–13. https://doi.org/10.1093/rheumatology/kep252
8. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315–24. https://doi.org/10.1002/art.1780310302
9. National Collaborating Centre for Chronic Conditions (UK). *Rheumatoid Arthritis: National Clinical Guideline for Management and Treatment in Adults*. NICE clinical guideline 79. London: Royal College of Physicians; 2009.
10. Combe B, Landewe R, Lukas C, Bolosiu HD, Breedveld F, Dougados M, *et al.* EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007;**66**:34–45. https://doi.org/10.1136/ard.2005.044354
11. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, *et al.* 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;**69**:1580–8. https://doi.org/10.1136/ard.2010.138461
12. Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, *et al.* The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology* 2002;**41**:793–800. https://doi.org/10.1093/rheumatology/41.7.793
13. Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Br J Rheumatol* 1994;**33**:735–9. https://doi.org/10.1093/rheumatology/33.8.735
14. Alamanos Y, Drosos AA. Epidemiology of adult rheumatoid arthritis. *Autoimmun Rev* 2005;**4**:130–6. https://doi.org/10.1016/j.autrev.2004.09.002

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

15. Goëb V, Fardellone P, Sibilia J, Ponchel F. Biomarkers in rheumatoid arthritis. *Mediators Inflamm* 2014;**2014**:379310. https://doi.org/10.1155/2014/379310

16. Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;**23**(5 Suppl. 39):S14–18.

17. Prevoo M, Van't Hof M, Kuper H, Van Leeuwen M, Van De Putte L, Van Riel P. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:44–8. https://doi.org/10.1002/art.1780380107

18. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;**23**:100–8.

19. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C *et al.* American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:727–35. https://doi.org/10.1002/art.1780380602

20. van Gestel AM, Prevoo ML, van't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996;**39**:34–40. https://doi.org/10.1002/art.1780390105

21. Felson DT. Assessing the efficacy and safety of rheumatic disease treatments: obstacles and proposed solutions. *Arthritis Rheum* 2003;**48**:1781–7. https://doi.org/10.1002/art.11087

22. van Gestel AM, Anderson JJ, van Riel PL, Boers M, Haagsma CJ, Rich B, *et al.* ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. American College of Rheumatology European League of Associations for Rheumatology. *J Rheumatol* 1999;**26**:705–11.

23. Stevenson M, Archer R, Tosh J, Simpson E, Everson-Hock E, Stevens JW, *et al.* Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying anti-rheumatic drugs and after the failure of conventional disease-modifying anti-rheumatic drugs only: systematic review and economic evaluation. *Health Technol Assess* 2016;**20**(35).

24. Wakefield RJ, D'Agostino MA, Naredo E, Buch MH, Iagnocco A, Terslev L, *et al.* After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? *Postgrad Med J* 2012;**88**:482–6. https://doi.org/10.1136/postgradmedj-2011-201048rep

25. Ben Abdelghani K, Miladi S, Souabni L, Kassab S, Chekili S, Laatar A, Zakraoui L. Role of ultrasound in assessing remission in rheumatoid arthritis. *Diagn Interv Imaging* 2015;**96**:3–10. https://doi.org/10.1016/j.diii.2014.07.006

26. Marks JL, Holroyd CR, Dimitrov BD, Armstrong RD, Calogeras A, Cooper C, *et al.* Does combined clinical and ultrasound assessment allow selection of individuals with rheumatoid arthritis for sustained reduction of anti-tumor necrosis factor therapy? *Arthritis Care Res (Hoboken)* 2015;**67**:746–53. https://doi.org/10.1002/acr.22552

27. Nishimoto N, Amano K, Hirabayashi Y, Horiuchi T, Ishii T, Iwahashi M, *et al.* Drug free REmission/low disease activity after cessation of tocilizumab (Actemra) Monotherapy (DREAM) study. *Mod Rheumatol* 2013;**24**:17–25. https://doi.org/10.3109/14397595.2013.854079

28. Colebatch AN, Edwards CJ, Østergaard M, van der Heijde D, Balint PV, D'Agostino MA, *et al.* EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis. *Ann Rheum Dis* 2013;**72**:804–14. https://doi.org/10.1136/annrheumdis-2012-203158

29. Aleo E, Barbieri F, Sconfienza L, Zampogna G, Garlaschi G, Cimmino MA. Ultrasound versus low-field magnetic resonance imaging in rheumatic diseases: a systematic literature review. *Clin Exp Rheum* 2014;**32**:91–8.

30. Taggart A, Benson C, Kane D. *Ultrasound in Rheumatology. Reports on the Rheumatic Diseases*, Series 6, Summer 2011, Topical Reviews No 9. 2011. URL: www.arthritisresearchuk.org/health-professionals-and-students/reports/reports-archives/~/media/Files/Education/Topical-Reviews/TR09-Summer-2011.ashx (accessed 25 January 2018).

31. van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 2000;**27**:261–3.

32. Genant HK, Jiang Y, Peterfy C, Lu Y, Redei J, Countryman PJ. Assessment of rheumatoid arthritis using a modified scoring method on digitized and original radiographs. *Arthritis Rheum* 1998;**41**:1583–90. https://doi.org/10.1002/1529-0131(199809)41:9<1583::AID-ART8>3.0.CO;2-H

33. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, *et al.* Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;**69**:631–7. https://doi.org/10.1136/ard.2009.123919

34. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P *et al.* Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;**75**:3–15. https://doi.org/10.1136/annrheumdis-2015-207524

35. Solomon DH, Bitton A, Katz JN, Radner H, Brown EM, Fraenkel L. *Review: treat to target in rheumatoid arthritis: fact, fiction, or hypothesis? Arthritis Rheumatol* 2014;**66**:775–82. https://doi.org/10.1002/art.38323

36. Wakefield RJ, D'Agostino MA, Naredo E, Buch MH, Iagnocco A, Terslev L, *et al.* After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? *Ann Rheum Dis* 2012;**71**:799–803. https://doi.org/10.1136/annrheumdis-2011-201048

37. Østergaard M, Møller-Bisgaard S. Rheumatoid arthritis: is imaging needed to define remission in rheumatoid arthritis? *Nat Rev Rheumatol* 2014;**10**:326–8. https://doi.org/10.1038/nrrheum.2014.63

38. Lugmani R, Hennell S, Estrach C, Birrell F, Bosworth A, Davenport G, *et al.* British Society for Rheumatology and British Health Professionals in Rheumatology Guideline for the Management of Rheumatoid Arthritis (the first 2 years). *Rheumatology (Oxford)* 2006;**45**:1167–9. https://doi.org/10.1093/rheumatology/kel215a

39. National Institute for Health and Care Excellence. *Rheumatoid Arthritis Overview*. URL: https://pathways.nice.org.uk/pathways/rheumatoid-arthritis#path=view%3A/pathways/rheumatoid-arthritis/managing-rheumatoid-arthritis.xml&content=view-index (accessed 31 March 2017).

40. National Institute for Health and Care Excellence (NICE). *Adalimumab, Etanercept, Infliximab, Certolizumab Pegol, Golimumab, Tocilizumab and Abatacept for Rheumatoid Arthritis Not Previously Treated with DMARDs or After Conventional DMARDs Only Have Failed*. Technology appraisal guidance TA375. London: NICE; 2016. URL: www.nice.org.uk/guidance/ta375 (accessed 25 January 2018).

41. National Institute for Health and Care Excellence (NICE). *Adalimumab, Etanercept, Infliximab, Rituximab and Abatacept for the Treatment of Rheumatoid Arthritis After the Failure of a TNF Inhibitor*. Technology appraisal guidance TA195. London: NICE; 2010. URL: www.nice.org.uk/guidance/ta195 (accessed 25 January 2018).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

42. National Institute for Health and Care Excellence (NICE). *Golimumab for the Treatment of Rheumatoid Arthritis After the Failure of Previous Disease-modifying Anti-rheumatic Drugs*. Technology appraisal guidance TA225. London: NICE; 2011. URL: www.nice.org.uk/guidance/ta225 (accessed 25 January 2018).

43. National Institute for Health and Care Excellence (NICE). *Tocilizumab for the Treatment of Rheumatoid Arthritis (Rapid Review of Technology Appraisal Guidance 198)*. Technology appraisal guidance TA247. London: NICE; 2012. URL: www.nice.org.uk/guidance/ta247 (accessed 25 January 2018).

44. National Institute for Health and Care Excellence (NICE). *Certolizumab Pegol for Treating Rheumatoid Arthritis after Inadequate Response to a TNF-Alpha Inhibitor*. Technology appraisal TA415. London: NICE; 2016. URL: www.nice.org.uk/guidance/ta415 (accessed 6 February 2018).

45. Mandl P, Naredo E, Wakefield RJ, Conaghan PG, D'Agostino MA. A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. *J Rheumatol* 2011;**38**:2055–62. https://doi.org/10.3899/jrheum.110424

46. McAlindon T, Kissin E, Nazarian L, Ranganath V, Prakash S, Taylor M, *et al.* American College of Rheumatology report on reasonable use of musculoskeletal ultrasonography in rheumatology clinical practice. *Arthritis Care Res* 2012;**64**:1625–40. https://doi.org/10.1002/acr.21836

47. Mulherin D, FitzGerald O, Bresnihan B. Clinical improvement and radiological deterioration in rheumatoid arthritis: evidence that the pathogenesis of synovial inflammation and articular erosion may differ. *Rheumatology (UK)* 1996;**35**:1263–8. https://doi.org/10.1093/rheumatology/35.12.1263

48. Molenaar ET, Voskuyl AE, Dinant HJ, Bezemer PD, Boers M, Dijkmans BA. Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. *Arthritis Rheum* 2004;**50**:36–42. https://doi.org/10.1002/art.11481

49. Cheung PP, Dougados M, Gossec L. Reliability of ultrasonography to detect synovitis in rheumatoid arthritis: a systematic literature review of 35 studies (1,415 patients). *Arthritis Care Res* 2010;**62**:323–34. https://doi.org/10.1002/acr.20102

50. Cheung PP, Ruyssen-Witrand A, Gossec L, Paternotte S, Le Bourlout C, Mazieres M, *et al.* Reliability of patient self-evaluation of swollen and tender joints in rheumatoid arthritis: a comparison study with ultrasonography, physician, and nurse assessments. *Arthritis Care Res* 2010;**62**:1112–19. https://doi.org/10.1002/acr.20178

51. Wakefield R. OMERACT 7 Special Interest Group. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005;**32**:2485–7.

52. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Østergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003;**48**:955–62. https://doi.org/10.1002/art.10877

53. Scheel AK, Hermann KG, Kahler E, Pasewaldt D, Fritz J, Hamm B, *et al.* A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. *Arthritis Rheum* 2005;**52**:733–43. https://doi.org/10.1002/art.20939

54. Naredo E, Bonilla G, Gamero F, Uson J, Carmona L, Laffon A. Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. *Ann Rheum Dis* 2005;**64**:375–81. https://doi.org/10.1136/ard.2004.023929

55. 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 Rheum* 2007;**57**:116–24. https://doi.org/10.1002/art.22461

56. Naredo E, Rodriguez M, Campos C, Rodriguez-Heredia JM, Medina JA, Giner E, *et al.* Validity, reproducibility, and responsiveness of a twelve-joint simplified power Doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. *Arthritis Rheum* 2008;**59**:515–22. https://doi.org/10.1002/art.23529

57. Karim Z, Wakefield RJ, Quinn M, Conaghan PG, Brown AK, Veale DJ, *et al.* Validation and reproducibility of ultrasonography in the detection of synovitis in the knee: a comparison with arthroscopy and clinical examination. *Arthritis Rheum* 2004;**50**:387–94. https://doi.org/10.1002/art.20054

58. Schmidt WA, Völker L, Zacher J, Schläfke M, Ruhnke M, Gromnica-Ihle E. Colour Doppler ultrasonography to detect pannus in knee joint synovitis. *Clin Exp Rheumatol* 2000;**18**:439–44.

59. Klauser A, Frauscher F, Schirmer M, Szkudlarek M, Court-Payen, Strandberg C, *et al.* Value of contrast-enhanced power Doppler ultrasonography (US) of the metacarpophalangeal joints on rheumatoid arthritis. *Eur Radiol* 2004;**14**:545–6.

60. D'Agostino MA, Wakefield R, Backhaus M, Balint PV, Brwyn GA, Filippucci E, *et al.* Combined evaluation of influence of sonographer and machine type on the reliability of power Doppler ultrasonography (PDUS) for detecting scoring and scanning synovitis in rheumatoid arthritis (RA) patients: results of an intermachine reliability exercise. *Ann Rheum Dis* 2008;**67**(Suppl. II):421.

61. Lazaar H, Lhoste-Trouilloud A, Pereira B, Couderc M, Mathieu S, Soubrier M. Does rheumatoid synovitis activity vary during the day? Evaluation with color Doppler sonography. *BMC Musculoskelet Disord* 2017;**18**:98. https://doi.org/10.1186/s12891-017-1450-3

62. D'Agostino MA, Wakefield RJ, Berner-Hammer H, Vittecoq O, Filippou G, Balint P, *et al.* Value of ultrasonography as a marker of early response to abatacept in patients with rheumatoid arthritis and an inadequate response to methotrexate: results from the APPRAISE study. *Ann Rheum Dis* 2016;**75**:1763–9. https://doi.org/10.1136/annrheumdis-2015-207709

63. Wakefield RJ, D'Agostino MA, Iagnocco A, Filippucci E, Backhaus M, Scheel AK, *et al.* The OMERACT Ultrasound Group: status of current activities and research directions. *J Rheumatol* 2007;**34**:848–51.

64. D'Agostino MA, Conaghan PG, Naredo E, Aegerter P, Iagnocco A, Freeston JE, *et al.* The OMERACT Ultrasound Task Force – advances and priorities. *J Rheumatol* 2009;**36**:1829–32. https://doi.org/10.3899/jrheum.090354

65. Iagnocco A, Naredo E, Wakefield R, Bruyn GA, Collado P, Jousse-Joulin S, *et al.* Responsiveness in rheumatoid arthritis. a report from the OMERACT 11 ultrasound workshop. *J Rheumatol* 2014;**41**:379–82. https://doi.org/10.3899/jrheum.131084

66. Naredo E, Rodriguez M, Rodriguez-Heredia J, Campos C, Medina J, Giner E, *et al.* Validity, reliability and sensitivity to change of simplified power Doppler ultrasonographic assessment of the effect of biological therapy on rheumatoid arthritis joint inflammation. *Ann Rheum Dis* 2007;**66**(Suppl. 2):183.

67. Backhaus M, Ohrndorf S, Kellner. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. *Arthritis Rheum* 2009;**61**:1194–201. https://doi.org/10.1002/art.24646

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

68. Hammer HB, Kvien TK. Comparisons of 7- to 78-joint ultrasonography scores: all different joint combinations show equal response to adalimumab treatment in patients with rheumatoid arthritis. *Arthritis Res Ther* 2011;**13**(Suppl. 3):R78. https://doi.org/10.1186/ar3341

69. 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. *Ann Rheum Dis* 2013;**72**:1163–9. https://doi.org/10.1136/annrheumdis-2012-201397

70. Naredo E, Wakefield RJ, Iagnocco A, Terslev L, Filippucci E, Gandjbakhch F, *et al.* The OMERACT ultrasound task force – status and perspectives. *J Rheumatol* 2011;**38**:2063–7. https://doi.org/10.3899/jrheum.110425

71. Cunnington J, Platt P, Raftery G, Kane D. Attitudes of United Kingdom rheumatologists to musculoskeletal ultrasound practice and training. *Ann Rheum Dis* 2007;**66**:1381–3. https://doi.org/10.1136/ard.2006.065466

72. Targeted Ultrasound Initiative. *About TUI*. URL: http://targetedultrasound.net/about (accessed 9 January 2018).

73. Department of Health and Social Care. *NHS Reference Costs 2014 to 2015*. 2016. URL: www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015 (accessed 27 October 2015).

74. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses:* the PRISMA statement. *BMJ* 2009;**339**:b2535. https://doi.org/10.1136/bmj.b2535

75. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, *et al.* Guidelines for musculoskeletal ultrasound in rheumatology. *Ann Rheum Dis* 2001;**60**:641–9. https://doi.org/10.1136/ard.60.7.641

76. Scottish Intercollegiate Guidelines Network (SIGN). *Management of Early Rheumatoid Arthritis*. SIGN publication no. 123. Edinburgh: SIGN; 2011. URL: www.sign.ac.uk/assets/sign123.pdf (accessed 20 October 2015).

77. National Institute for Health and Clinical Excellence, Centre for Clinical Practice – Surveillance Programme. *Surveillance Report for Guidance Executive (Update CG79)*. 2015. URL: www.nice.org.uk/guidance/CG79/update/CG79/documents/cg79-rheumatoid-arthritis-surveillance-review-decision-march-20153 (accessed 27 October 2015).

78. Dale J, Stirling A, Zhang R, Purves D, Foley J, Sambrook M, *et al.* Targeting ultrasound remission in early rheumatoid arthritis: the results of the TaSER study, a randomised clinical trial. *Ann Rheum Dis* 2016;**75**:1043–50. https://doi.org/10.1136/annrheumdis-2015-208941

79. Haavardsholm E, Aga A-B, Olsen I, Hammer HB, Uhlig T, Fremstad H, *et al.* Aiming for remission in rheumatoid arthritis: clinical and radiographic outcomes from a randomized controlled strategy trial investigating the added value of ultrasonography in a treat-to-target regimen. *Arthritis Rheumatol* 2015;**67**(Suppl. 10):10L.

80. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000;**53**:459–68. https://doi.org/10.1016/S0895-4356(99)00206-1

81. Williams K, Moons C, for the Cochrane Prognosis Review Methods Group. *An Introduction to Systematic Reviews of Prognosis*. Auckland, New Zealand: Cochrane Prognosis Review Methods Group; 2012.

82. Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. Wellington, New Zealand: *BMC Med Res Methodol* 2003;**3**:25. https://doi.org/10.1186/1471-2288-3-25

83. New Zealand Guidelines Group. *Handbook for Preparation of Explicit Evidence-Based Clinical Practice Guidelines*. Wellington: New Zealand Guidelines Group; 2001.

84. Karsh J, Keystone EC, Haraoui B, Thorne JC, Pope JE, Bykerk VP, *et al.* Canadian recommendations for clinical trials of pharmacologic interventions in rheumatoid arthritis: inclusion criteria and study design. *J Rheumatol* 2011;**38**:2095–104. https://doi.org/10.3899/jrheum.110188

85. Alcalde M, D'Agostino MA, Bruyn GAW, Moller I, Iagnocco A, Wakefield RJ, *et al.* A systematic literature review of US definitions, scoring systems and validity according to the OMERACT filter for tendon lesion in RA and other inflammatory joint diseases. *Rheumatology (Oxford)* 2012;**51**:1246–60. https://doi.org/10.1093/rheumatology/kes018

86. Joshua F, Lassere M, Bruyn GA, Szkudlarek M, Naredo E, Schmidt WA, *et al.* Summary findings of a systematic review of the ultrasound assessment of synovitis. *J Rheumatol* 2007;**34**:839–47.

87. Keen HI, Mease PJ, Bingham CO, Giles JT, Kaeley G, Conaghan PG. Systematic review of MRI, ultrasound, and scintigraphy as outcome measures for structural pathology in interventional therapeutic studies of knee arthritis: focus on responsiveness. *J Rheumatol* 2011;**38**:142–54. https://doi.org/10.3899/jrheum.100377

88. Ten Cate DF, Luime JJ, Swen N, Gerards AH, De Jager MH, Basoski NM. Role of ultrasonography in diagnosing early rheumatoid arthritis and remission of rheumatoid arthritis – a systematic review of the literature. *Arthritis Res Ther* 2013;**15**(Suppl. 1):R4. https://doi.org/10.1186/ar4132

89. Bhamra K, Seymour M, Swales C, Taylor P. Utility of ultrasound in the Nuffield Orthopaedic Centre Emergency Rheumatology Clinic: survey of clinical effectiveness. *Ann Rheum Dis* 2014;**73**:659–60. https://doi.org/10.1136/annrheumdis-2014-eular.4650

90. Ciurtin C, Ehrenstein M, Leandro M, Dey D, Nandagudi A, Morris V. *et al.* The usefulness of a musculoskeletal ultrasound (MUS) scoring system for 22 hand joints examination for the detection of early undifferentiated inflammatory arthritis and treatment decisions making in established inflammatory arthritis. *Ann Rheum Dis* 2013;**71**:295–6. https://doi.org/10.1136/annrheumdis-2012-eular.2381

91. Gandjbakhch F, Millot F, Etchepare F, Foltz V, Bourgeois P, Fautrel B. Ultrasonography examination influences therapeutic decision in RA patients. *Arthritis Rheum* 2008;**58**(9 Suppl. S):S467–68.

92. Hayashi N, Ogura T, Hirata A, Kujime R, Imamura M, Takenaka S, *et al.* The comparison between physical and ultrasound joint examination for the hand joints in patients with rheumatoid arthritis. *Arthritis Rheumatol* 2014;**66**:S181.

93. Inanc N, Ozen G, Direskeneli H. Ultrasonographic assessment of joint inflammation in rheumatoid arthritis: predictive value in response to tumor necrosis factor-alpha inhibitor treatment. *Ann Rheum Dis* 2014;**73**:468. https://doi.org/10.1136/annrheumdis-2014-eular.1918

94. Kelly S, Davidson B, Gorman C, Meenagh G, Reynolds P. The impact of ultrasound on the diagnosis and management of patients with rheumatoid arthritis (RA) in routine clinical care within the UK. *Ann Rheum Dis* 2013;**72**:A101. https://doi.org/10.1136/annrheumdis-2013-eular.351

95. Luengroongroj P, Suwannalai P, Joavisidha S. Disease flared after DMARD tapering in rheumatoid arthritis may be predicted by subclinical synovitis detected by ultrasonography. *Int J Rheum Dis* 2015;**18**(Suppl. S1):107.

96. Mamoto K, Koike T, Okano T, Tada M, Sugioka Y, Wakitani S, *et al.* Comparative assessment of swollen joints in rheumatoid arthritis by patients, physicians and ultrasonography. *Ann Rheum Dis* 2013;**71**:710 https://doi.org/10.1136/annrheumdis-2012-eular.1272

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

97. 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. *Ann Rheum Dis* 2013;**72**(Suppl. 3):A755–6. https://doi.org/10.1136/annrheumdis-2013-eular.2238

98. Ramirez García J, Ruíz-Esquide V, Celis R, Cuervo A, Cabrera S, Inciarte-Mundo J, *et al.* Sonographic and clinical characterization of a prospective cohort of patients with rheumatoid arthritis in clinical remission. Preliminary results. *Ann Rheum Dis* 2014;**73**:888 https://doi.org/10.1136/annrheumdis-2014-eular.4310

99. Zufferey P. Persistence of ultrasound synovitis in patients with rheumatoid arthritis fulfilling the DAS28 and/or the new ACR/EULAR RA remission definitions: results of an observational cohort study. *Joint Bone Spine* 2014;**81**:426–32. https://doi.org/10.1016/j.jbspin.2014.04.014

100. Taylor PC, Steuer A, Gruber J, Cosgrove DO, Blomley MJK, Marsters A, *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 Rheum* 2004;**50**:1107–16. https://doi.org/10.1002/art.20123

101. Ellegaard K, Christensen R, Torp-Pedersen S, Terslev L, Holm CC, Konig MJ, *et al.* Ultrasound Doppler measurements predict success of treatment with anti-TNF-alpha drug in patients with rheumatoid arthritis: a prospective cohort study. *Rheumatology (Oxford)* 2011;**50**:506–12. https://doi.org/10.1093/rheumatology/keq336

102. Kane D, Balint PV, Sturrock RD. Ultrasonography is superior to clinical examination in the detection and localization of knee joint effusion in rheumatoid arthritis. *J Rheumatol* 2003;**30**:966–71.

103. Luukkainen R, Sanila MT. Relationship between clinically detected joint swelling and effusion diagnosed by ultrasonography in elbow joints in patients with rheumatoid arthritis. *Clin Rheumatol* 2005;**24**:228–31. https://doi.org/10.1007/s10067-004-1010-8

104. Luukkainen R, Sanila MT. Poor relationship between joint swelling detected on physical examination and effusion diagnosed by ultrasonography in glenohumeral joints in patients with rheumatoid arthritis. *Clin Rheumatol* 2007;**26**:865–7. https://doi.org/10.1007/s10067-006-0402-3

105. Balsa A, de Miguel E, Castillo C, Peiteado D, Martin-Mola E. Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. *Rheumatology* 2010;**49**:683–90. https://doi.org/10.1093/rheumatology/kep442

106. Beckers C, Ribbens C, André B, Marcelis S, Kaye O, Mathy L, *et al.* Assessment of disease activity in rheumatoid arthritis with (18)F-FDG PET. *J Nucl Med* 2004;**45**:956–64.

107. Filippucci E, Iagnocco A, Salaffi F, Cerioni A, Valesini G, Grassi W. Power Doppler sonography monitoring of synovial perfusion at the wrist joints in patients with rheumatoid arthritis treated with adalimumab. *Ann Rheum Dis* 2006;**65**:1433–7. https://doi.org/10.1136/ard.2005.044628

108. Garrigues F, Jousse-Joulin S, Bouttier R, Nonent M, Bressollette L, Saraux A. Concordance between clinical and ultrasound findings in rheumatoid arthritis. *Joint Bone Spine* 2013;**80**:597–603. https://doi.org/10.1016/j.jbspin.2013.03.011

109. Gartner M, Mandl P, Radner H, Supp G, Machold KP, Aletaha D, *et al.* Sonographic joint assessment in rheumatoid arthritis: associations with clinical joint assessment during a state of remission. *Arthritis Rheum* 2013;**65**:2005–14. https://doi.org/10.1002/art.38016

110. Haavardsholm EA, Ostergaard M. Monitoring anti-TNFalpha treatment in rheumatoid arthritis: responsiveness of magnetic resonance imaging and ultrasonography of the dominant wrist joint compared with conventional measures of disease activity and structural damage. *Ann Rheum Dis* 2009;**68**:1572–9. https://doi.org/10.1136/ard.2008.091801

111. Hammer HB, Sveinsson M, Kongtorp AK, Kvien TK. A 78-joints ultrasonographic assessment is associated with clinical assessments and is highly responsive to improvement in a longitudinal study of patients with rheumatoid arthritis starting adalimumab treatment. *Ann Rheum Dis* 2010;**69**:1349–51. https://doi.org/10.1136/ard.2009.126995

112. Horikoshi M, Suzuki T, Sugihara M, Kondo Y, Tsuboi H, Uehara T, *et al.* Comparison of low-field dedicated extremity magnetic resonance imaging with articular ultrasonography in patients with rheumatoid arthritis. *Mod Rheumatol* 2010;**20**:556–60. https://doi.org/10.3109/s10165-010-0318-2

113. 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. *J Rheumatol* 2013;**40**:1967–76. https://doi.org/10.3899/jrheum.130556

114. Kamishima T, Tanimura K, Shimizu M, Matsuhashi M, Fukae J, Kon Y, *et al.* Monitoring anti-interleukin 6 receptor antibody treatment for rheumatoid arthritis by quantitative magnetic resonance imaging of the hand and power Doppler ultrasonography of the finger. *Skeletal Radiol* 2011;**40**:745–55. https://doi.org/10.1007/s00256-010-1064-4

115. Luukkainen RK, Saltyshev M. Relationship between clinically detected joint swelling and effusion diagnosed by ultrasonography in metatarsophalangeal and talocrural joints in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2003;**21**:632–4.

116. Mandl P, Balint PV. Clinical and ultrasound-based composite disease activity indices in rheumatoid arthritis: results from a multicenter, randomized study. *Arthritis Care Res (Hoboken)* 2013;**65**:879–87. https://doi.org/10.1002/acr.21913

117. Mandl P, Balint PV. Metrologic properties of ultrasound versus clinical evaluation of synovitis in rheumatoid arthritis: results of a multicenter, randomized study. *Arthritis Rheum* 2012;**64**:1272–82. https://doi.org/10.1002/art.33491

118. Naredo E, Valor L, De la Torre I, Martinez-Barrio J, Hinojosa M, Aramburu F, *et al.* Ultrasound joint inflammation in rheumatoid arthritis in clinical remission: how many and which joints should be assessed? *Arthritis Care Res (Hoboken)* 2013;**65**:512–17. https://doi.org/10.1002/acr.21869

119. Pereira DF, Gutierrez M, de Buosi AL, Ferreira FB, Draghessi A, Grassi W, *et al.* Is articular pain in rheumatoid arthritis correlated with ultrasound power Doppler findings? *Clin Rheumatol* 2015;**34**:1975–9. https://doi.org/10.1007/s10067-015-2964-4

120. Ribbens C, Andre B, Marcelis S, Kaye O, Mathy L, Bonnet V, *et al.* Rheumatoid hand joint synovitis: gray-scale and power Doppler US quantifications following anti-tumor necrosis factor-alpha treatment: pilot study. *Radiology* 2003;**229**:562–9. https://doi.org/10.1148/radiol.2292020206

121. Riente L, Delle Sedie A, Filippucci E, Scirè CA, Iagnocco A, Gutierrez M, *et al.* Ultrasound imaging for the rheumatologist. XXVII. Sonographic assessment of the knee in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2010;**28**:300–3.

122. Riente L, Delle Sedie A, Scirè CA, Filippucci E, Meenagh G, Iagnocco A, *et al.* Ultrasound imaging for the rheumatologist. XXXI. Sonographic assessment of the foot in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2011;**29**:1–5.

123. Salaffi F, Filippucci E, Carotti M, Naredo E, Meenagh G, Ciapetti A, *et al.* Interobserver agreement of standard joint counts in early rheumatoid arthritis: a comparison with grey-scale ultrasonograpky. *Ann Rheum Dis* 2006;**65**:592–3.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

124. Saleem B, Brown AK. Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. *Ann Rheum Dis* 2011;**70**:792–8. https://doi.org/10.1136/ard.2010.134445

125. Spiegel TM, King W III, Weiner SR, Paulus HE. Measuring disease activity: comparison of joint tenderness, swelling, and ultrasonography in rheumatoid arthritis. *Arthritis Rheum* 1987;**30**:1283–8. https://doi.org/10.1002/art.1780301111

126. Szkudlarek M, Narvestad E, Klarlund M, Court-Payen, Thomsen HS, Ostergaard M, *et al.* Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography, and clinical examination. *Arthritis Rheum* 2004;**50**:2103–12. https://doi.org/10.1002/art.20333

127. Szkudlarek M, Klarlund M, Narvestad E, Court-Payen, Strandberg C, Jensen KE, *et al.* Ultrasonography of the metacarpophalangeal and proximal interphalangeal joints in rheumatoid arthritis: a comparison with magnetic resonance imaging, conventional radiography and clinical examination. *Arthritis Res Ther* 2006;**8**(Suppl. 2):R52. https://doi.org/10.1186/ar1904

128. Taniguchi D, Tokunaga D, Oda R, Fujiwara H, Ikeda T, Ikoma K, *et al.* Maximum intensity projection with magnetic resonance imaging for evaluating synovitis of the hand in rheumatoid arthritis: comparison with clinical and ultrasound findings. *Clin Rheumatol* 2014;**33**:911–17. https://doi.org/10.1007/s10067-014-2526-1

129. 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. *Med Ultrason* 2015;**17**:352–60. https://doi.org/10.11152/mu.2013.2066.173.viv

130. Wakefield RJ, Freeston JE, O'Connor P, Reay N, Budgen A, Hensor EM, *et al.* The optimal assessment of the rheumatoid arthritis hindfoot: a comparative study of clinical examination, ultrasound and high field MRI. *Ann Rheum Dis* 2008;**67**:1678–82. https://doi.org/10.1136/ard.2007.079947

131. Xiao H, Liu M, Tan L, Liao X, Li Y, Gao J, *et al.* Value of ultrasonography for diagnosis of synovitis associated with rheumatoid arthritis. *Int J Rheum Dis* 2014;**17**:767–75. https://doi.org/10.1111/1756-185X.12390

132. Zufferey P. Ultrasound evaluation of synovitis in RA: correlation with clinical disease activity and sensitivity to change in an observational cohort study. *Joint Bone Spine* 2014;**81**:222–7. https://doi.org/10.1016/j.jbspin.2013.08.006

133. Gartner M, Radner H, Supp G, Mandl P, Machold K, Aletaha D, *et al.* Does joint sonography really add clinically important information beyond clinical joint examination? *J Miner Stoffwechs* 2012;**19**(Suppl. 4):174.

134. Boyesen P, Haavardsholm EA. Prediction of MRI erosive progression: a comparison of modern imaging modalities in early rheumatoid arthritis patients. *Ann Rheum Dis* 2011;**70**:176–9. https://doi.org/10.1136/ard.2009.126953

135. Brown AK, Conaghan PG, Karim, Conaghan PG. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008;**58**:2958–67. https://doi.org/10.1002/art.23945

136. Ikeda K, Brown A, Conaghan P, Karim Z, Quinn M, Peterfy C, *et al.* An explanation of structural deterioration in RA patients in clinical remission on DMARDS: the significance of sub-clinical inflammation detectable with imaging. *Ann Rheum Dis* 2007;**66**(Suppl. 2):95.

137. Scirè 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 (Oxford)* 2009;**48**:1092–7. https://doi.org/10.1093/rheumatology/kep171

138. Cavet G, Shen Y, Abraham S, Chernoff D, Centola M, Taylor P. Predicting radiographic progression in rheumatoid arthritis with ultrasound and biomarkers. *Arthritis Rheum* 2009;**60**:1464.

139. Dougados M, Devauchelle-Pensec V, Ferlet JF, Jousse-Joulin S, D'Agostino M-A, Backhaus M, *et al.* The ability of synovitis to predict structural damage in rheumatoid arthritis: a comparative study between clinical examination and ultrasound. *Ann Rheum Dis* 2013;**72**:665–71. [Erratum published in *Ann Rheum Dis* 2013;**72**:1270.] https://doi.org/10.1136/annrheumdis-2012-201469

140. 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 Rheum* 2008;**58**:2248–56. https://doi.org/10.1002/art.23682

141. 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 Radiol* 2009;**38**:473–8. https://doi.org/10.1007/s00256-009-0670-5

142. 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. *Ann Rheum Dis* 2012;**71**:1316–21. https://doi.org/10.1136/annrheumdis-2011-200548

143. Wakefield RJ, Freeston JE, Hensor EMA, 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 Rheum* 2007;**57**:1564–7. https://doi.org/10.1002/art.23097

144. 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. *Mod Rheumatol* 2013;**23**:456–65. https://doi.org/10.3109/s10165-012-0690-1

145. 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. *Mod Rheumatol* 2014;**24**:590–8. https://doi.org/10.3109/14397595.2013.857800

146. Ikeda K, Nakagomi D, Sanayama Y, Yamagata M, Okubo A, Iwamoto T, *et al.* 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 Rheum* 2012;**64**:S50.

147. Bugatti S, Manzo, Manzo A, Benaglio F, Klersy C, Vitolo B, *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 Res Ther* 2012;**14**(Suppl. 1):R34. https://doi.org/10.1186/ar3742

148. 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 Rheumatol* 2014;**66**(Suppl. 10):S602.

149. Dougados M, Devauchelle-Pensec V, Ferlet J-F, Jousse-Joulin S, D'Agostino M-A, Backhaus M, *et al.* Both clinical and ultrasonographic evaluation of synovitis are relevant to predict subsequent radiological deterioration in rheumatoid arthritis. *Ann Rheum Dis* 2013;**71**:610. https://doi.org/10.1136/annrheumdis-2012-eular.3354

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

150. 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. *Ann Rheum Dis* 2013;**72**:A755. https://doi.org/10.1136/annrheumdis-2013-eular.2237

151. 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 (Oxford)* 2007;**46**:454–9. https://doi.org/10.1093/rheumatology/kel256

152. Ceponis A, Onishi M, Bluestein HG, Kalunian K, Townsend J, Kavanaugh A. Utility of the ultrasound examination of the hand and wrist joints in the management of established rheumatoid arthritis. *Arthritis Care Res* 2014;**66**:236–44. https://doi.org/10.1002/acr.22119

153. 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 Res* 2014;**66**:19–26. https://doi.org/10.1002/acr.22218

154. Dale J, Stirling A, McInnes IB, Porter D. Targeting ultrasound remission in early rheumatoid arthritis – results of the TaSER study. *Arthritis Rheum* 2013;**65**:S338–9.

155. Iwamoto T, Iwamoto T, Ikeda K, Hosokawa J, Yamagata M, Tanaka S, *et al.* Prediction of relapse after discontinuation of biologic agents by ultrasonographic assessment in patients with rheumatoid arthritis in clinical remission: high predictive values of total gray-scale and power Doppler scores that represent residual synovial inflammation before discontinuation. *Arthritis Care Res (Hoboken)* 2014;**66**:1576–81. https://doi.org/10.1002/acr.22303

156. Naredo E, Valor L, De la Torre, I, Montoro M, Bello N, Martinez-Barrio J, *et al.* Predictive value of Doppler ultrasound-detected synovitis in relation to successful tapering of biology therapy in patients with rheumatoid arthritis. *Ann Rheum Dis* 2014;**73**:269. https://doi.org/10.1136/annrheumdis-2014-eular.3013

157. Naredo E, Valor L, De la Torre, I, Montoro M, Bello N, Martinez-Barrio J, *et al.* Predictive value of Doppler ultrasound-detected synovitis in relation to failed tapering of biologic therapy in patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2015;**54**:1408–14. https://doi.org/10.1093/rheumatology/kev006

158. Ciurtin C, Leandro M, Dey D, Nandagudi A, Giles I, Shipley M, *et al.* The usefulness of a musculoskeletal ultrasound (MUS) scoring system for 22 hand joints examination for the detection of early undifferentiated inflammatory arthritis and treatment decision making in established inflammatory arthritis. *Rheumatology (Oxford)* 2012;**51**:iii70.

159. Merlin T, Weston A, Tooher R. Extending an evidence hierarchy to include topics other than treatment: revising the Australian 'levels of evidence'. *BMC Med Res Methodol* 2009;**9**:34. https://doi.org/10.1186/1471-2288-9-34

160. Salaffi F, Filippucci E, Carotti. Inter-observer agreement of standard joint counts in early rheumatoid arthritis: a comparison with grey scale ultrasonography – a preliminary study. *Rheumatology* 2008;**47**:54–8. https://doi.org/10.1093/rheumatology/kem286

161. Dale J, Purves D, McConnachie A, Porter D, McInnes IB. Tightening up: musculoskeletal ultrasound could further individualise treatment decisions in early rheumatoid arthritis patients treated by a step-up DMARD escalation regimen. *Arthritis Rheum* 2012;**64**:S1129.

162. ClinicalTrials.gov. *Aiming for Remission in Rheumatoid Arthritis (RA) – the ARCTIC Trial (ARCTIC)*. 2016. URL: https://clinicaltrials.gov/ct2/show/study/NCT01205854?term=NCT+01205854&rank=1&view=record (accessed 20 October 2015).

163. Holroyd CR, Davidson B, Bennett S, Waghorn D, Underhil C, Cooper C, *et al.* A strategy for selecting individuals with RA for reduction of anti-TNF therapy using combined clinical and ultrasound assessment. *Arthritis Rheum* 2013;**65**:S339–40.

164. Lamers-Karnebeek FBG, Jansen T, van Riel P, Luime J, Jacobs J. Ultrasonography as predictor for flare in rheumatoid arthritis patients with low disease activity: nine month results from POET-US-study. *Ann Rheum Dis* 2015;**74**:140. https://doi.org/10.1136/annrheumdis-2015-eular.5185

165. Arshad A, Sulaiman W. Optimizing the use of traditional DMARD in RA: getting the most out of what we can afford! *J Rheumatol* 2007;**10**:3–8. https://doi.org/10.1111/j.1479-8077.2007.00247.x

166. Garg OP, Bhakuni DS, Narayanan K, Vasdev V, Jain R. Guided interventions in rheumatology: Our experiences and innovations at a premier institute. *Int J Rheum Dis* 2010;**13**:223.

167. Gibson N, Kissin E. The pros and cons of ultrasonography for rheumatologic conditions. *J Musculoskelet Med* 2011;**28**:289–95.

168. Østergaard M, Døhn UM, Ejbjerg BJ, McQueen FM. Ultrasonography and magnetic resonance imaging in early rheumatoid arthritis: recent advances. *Curr Rheumatol Rep* 2006;**8**:378–85. https://doi.org/10.1007/s11926-006-0069-4

169. van Ingen IL, Lamers-Karnebeek F, Jansen TL. Optimizing the expediency of TNFi in rheumatoid arthritis: offering a TNFi holiday in patients having reached low-disease activity in the maintenance phase. *Expert Opin Biol Ther* 2014;**14**:1761–7. https://doi.org/10.1517/14712598.2014.955009

170. Breedveld F. The value of early intervention in RA – a window of opportunity. *Clin Rheumatol* 2011;**30**:33–9. https://doi.org/10.1007/s10067-010-1638-5

171. Brown PM, Isaacs JD. Rheumatoid arthritis: an evolutionary force in biologics. *Curr Pharm Des* 2015;**21**:2170–8. https://doi.org/10.2174/1381612821666150310141827

172. Caporali R, Scirè CA, Todoerti M, Montecucco C. The role of low-dose glucocorticoids for rheumatoid arthritis in the biologic era. *Clin Exp Rheumatol* 2013;**31**:9–13.

173. Claessen SJ, Hazes JM, Huisman MA, Van ZD, Luime JJ, Weel AE. Use of risk stratification to target therapies in patients with recent onset arthritis; design of a prospective randomized multicenter controlled trial. *BMC Musculoskelet Disord* 2009;**10**:71. https://doi.org/10.1186/1471-2474-10-71

174. Favalli EG, Marchesoni A, Colombo GL, Sinigaglia L. Pattern of use, economic burden and vial optimization of infliximab for rheumatoid arthritis in Italy. *Clin Exp Rheumatol* 2008;**26**:45–51.

175. Heather EM, Payne K, Harrison M, Symmons DP. Including adverse drug events in economic evaluations of anti-tumour necrosis factor-alpha drugs for adult rheumatoid arthritis: a systematic review of economic decision analytic models. *Pharmacoeconomics* 2014;**32**:109–34. https://doi.org/10.1007/s40273-013-0120-z

176. Kobelt G, Lekander I, Lang A, Raffeiner B, Botsios C, Geborek P. Cost-effectiveness of etanercept treatment in early active rheumatoid arthritis followed by dose adjustment. *Int J Technol Assess Health Care* 2011;**27**:193–200. https://doi.org/10.1017/S0266462311000195

177. Kobelt G. Treating to target with etanercept in rheumatoid arthritis: cost-effectiveness of dose reductions when remission is achieved. *Value Health* 2014;**17**:537–44. https://doi.org/10.1016/j.jval.2014.04.005

178. Krieckaert CL, Nair SC, Nurmohamed MT, van Dongen CJ, Lems WF, Lafeber FP, *et al.* Personalised treatment using serum drug levels of adalimumab in patients with rheumatoid arthritis: an evaluation of costs and effects. *Ann Rheum Dis* 2015;**74**:361–8. https://doi.org/10.1136/annrheumdis-2013-204101

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

179. Radner H, Aletaha D. Anti-TNF in rheumatoid arthritis: an overview. *Wien Med Wochenschr* 2015;**165**:3–9. https://doi.org/10.1007/s10354-015-0344-y

180. Saleem B, Nizam S, Emery P. Can remission be maintained with or without further drug therapy in rheumatoid arthritis? *Clin Exp Rheumatol* 2006;**24**:33–6.

181. Scott IC, Wailoo A, Scott DL. Payers' views on treating-to-target in rheumatoid arthritis: an English perspective. *Clin Exp Rheumatol* 2012;**30**(Suppl. 4):85–90.

182. Smolen JS, Emery P, Fleischmann R, Van Vollenhoven RF, Pavelka K, Durez P, *et al.* Adjustment of therapy in rheumatoid arthritis on the basis of achievement of stable low disease activity with adalimumab plus methotrexate or methotrexate alone: the randomised controlled OPTIMA trial. *Lancet* 2014;**383**:321–32. [Erratum published in Lancet 2014;**383**:308.] https://doi.org/10.1016/S0140-6736(13)61751-1

183. Smolen JS, Nash P, Durez P, Hall S, Ilivanova E, Irazoque-Palazuelos F, *et al.* Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. *Lancet* 2013;**381**:918–29. https://doi.org/10.1016/S0140-6736(12)61811-X

184. Tanaka Y, Takeuchi T, Mimori T, Saito K, Nawata M, Kameda H, *et al.* Discontinuation of infliximab after attaining low disease activity in patients with rheumatoid arthritis: RRR (remission induction by Remicade in RA) study. *Ann Rheum Dis* 2010;**69**:1286–91. https://doi.org/10.1136/ard.2009.121491

185. Tanaka Y, Hirata S. Is it possible to withdraw biologics from therapy in rheumatoid arthritis? *Clin Ther* 2013;**35**:2028–35. https://doi.org/10.1016/j.clinthera.2013.10.008

186. Tanaka Y. Next stage of RA treatment: is TNF inhibitor-free remission a possible treatment goal? *Ann Rheum Dis* 2013;**72**(Suppl. 2):ii124–7. https://doi.org/10.1136/annrheumdis-2012-202350

187. van Herwaarden N, van der Maas A, Minten MJ, van den Hoogen FH, Kievit W, Van Vollenhoven RF, *et al.* Disease activity guided dose reduction and withdrawal of adalimumab or etanercept compared with usual care in rheumatoid arthritis: open label, randomised controlled, non-inferiority trial. *BMJ* 2015;**350**:h1389. https://doi.org/10.1136/bmj.h1389

188. Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J *et al.* Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). *Ann Rheum Dis* 2016;**75**:59–67. https://doi.org/10.1136/annrheumdis-2014-206696

189. Galloway JB, Kingsley G, Ma M, Lorente-Canovas B, Cope A, Ibrahim F, *et al.* SAT0150 Optimising Treatment with TNF inhibitors in rheumatoid arthritis with different dose tapering strategies: the Opttira trial. *Ann Rheum Dis* 2015;**74**:706. https://doi.org/10.1136/annrheumdis-2015-eular.4684

190. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, van Riel PL, van de Laar MA, Jansen TL. SAT0157 Randomized Trial of Stopping TNF-Inhibitors in Rheumatoid Arthritis Patients with Stable Remission or Low Disease Activity in the Netherlands. *Ann Rheum Dis* 2015;**74**(Suppl. 2):709. https://doi.org/10.1136/annrheumdis-2015-eular.2388

191. Tanaka Y, Hirata S, Kubo S, Fukuyo S, Hanami K, Sawamukai N, *et al.* Discontinuation of adalimumab after achieving remission in patients with established rheumatoid arthritis: 1-year outcome of the HONOR study. *Ann Rheum Dis* 2015;**74**:389–5. https://doi.org/10.1136/annrheumdis-2013-204016

192. van Vollenhoven RF, Østergaard M, Leirisalo-Repo M, Uhlig T, Jansson M, Larsson E, *et al.* Full dose, reduced dose or discontinuation of etanercept in rheumatoid arthritis. *Ann Rheum Dis* 2016;**75**:52–8. https://doi.org/10.1136/annrheumdis-2014-205726

193. Emery P, Hammoudeh M, FitzGerald O, Combe B, Martin-Mola E, Buch MH, *et al.* Sustained remission with etanercept tapering in early rheumatoid arthritis. *N Engl J Med* 2014;**371**:1781–92. https://doi.org/10.1056/NEJMoa1316133

194. Kavanaugh A, Fleischmann RM, Emery P, Kupper H, Redden L, Guerette B, *et al.* Clinical, functional and radiographic consequences of achieving stable low disease activity and remission with adalimumab plus methotrexate or methotrexate alone in early rheumatoid arthritis: 26-week results from the randomised, controlled OPTIMA study. *Ann Rheum Dis* 2013;**72**:64–71. https://doi.org/10.1136/annrheumdis-2011-201247

195. Aguilar-Lozano L, Castillo-Ortiz JD, Vargas-Serafin C, Morales-Torres J, Sanchez-Ortiz A, Sandoval-Castro C, *et al.* Sustained clinical remission and rate of relapse after tocilizumab withdrawal in patients with rheumatoid arthritis. *J Rheumatol* 2013;**40**:1069–73. https://doi.org/10.3899/jrheum.121427

196. Detert J, Bastian H, Listing J, Weiss A, Wassenberg S, Liebhaber A, *et al.* Induction therapy with adalimumab plus methotrexate for 24 weeks followed by methotrexate monotherapy up to week 48 versus methotrexate therapy alone for DMARD-naive patients with early rheumatoid arthritis: HIT HARD, an investigator-initiated study. *Ann Rheum Dis* 2013;**72**:844–50. https://doi.org/10.1136/annrheumdis-2012-201612

197. Smolen JS, Emery P, Ferraccioli GF, Samborski W, Berenbaum F, Davies OR, *et al.* Certolizumab pegol in rheumatoid arthritis patients with low to moderate activity: the CERTAIN double-blind, randomised, placebo-controlled trial. *Ann Rheum Dis* 2014;**74**:843–50. https://doi.org/10.1136/annrheumdis-2013-204632

198. Chatzidionysiou K, Turesson C, Teleman A, Knight A, Lindqvist E, Larsson P, *et al.* A multicenter, randomized, controlled, open-label pilot study of the feasibility of discontinuation of adalimumab in rheumatoid arthritis patients in stable clinical remission. *RMD Open* 2016;**2**:e000133. https://doi.org/10.1136/rmdopen-2015-000133

199. Harigai M, Takeuchi T, Tanaka Y, Matsubara T, Yamanaka H, Miyasaka N. Discontinuation of adalimumab treatment in rheumatoid arthritis patients after achieving low disease activity. *Mod Rheumatol* 2012;**22**:814–22. https://doi.org/10.3109/s10165-011-0586-5

200. Kaine J, Gladstein G, Strusberg I, Robles M, Louw I, Gujrathi S, *et al.* Evaluation of abatacept administered subcutaneously in adults with active rheumatoid arthritis: impact of withdrawal and reintroduction on immunogenicity, efficacy and safety (phase IIIb ALLOW study). *Ann Rheum Dis* 2012;**71**:38–44. https://doi.org/10.1136/annrheumdis-2011-200344

201. van der Maas A, Kievit W, van den Bemt BJ, van den Hoogen FH, van Riel PL, den Broeder AA. Down-titration and discontinuation of infliximab in rheumatoid arthritis patients with stable low disease activity and stable treatment: an observational cohort study. *Ann Rheum Dis* 2012;**71**:1849–54. https://doi.org/10.1136/annrheumdis-2011-200945

202. Klarenbeek NB, van der Kooij SM, Guler-Yuksel M, van Groenendael JH, Han KH, Kerstens PJ, *et al.* Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. *Ann Rheum Dis* 2011;**70**:315–19. https://doi.org/10.1136/ard.2010.136556

203. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, Van Zeben D, Kerstens PJ, Hazes JM, *et al.* Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. *Arthritis Rheum* 2005;**52**:3381–90. https://doi.org/10.1002/art.21405

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

204. van den Broek M, Klarenbeek NB, Dirven L, Van Schaardenburg D, Hulsmans HM, Kerstens PJ, *et al.* Discontinuation of infliximab and potential predictors of persistent low disease activity in patients with early rheumatoid arthritis and disease activity score-steered therapy: subanalysis of the BeSt study. *Ann Rheum Dis* 2011;**70**:1389–94. https://doi.org/10.1136/ard.2010.147751

205. Bejarano V, Conaghan PG, Quinn MA, Saleem B, Emery P. Benefits 8 years after a remission induction regime with an infliximab and methotrexate combination in early rheumatoid arthritis. *Rheumatology (Oxford)* 2010;**49**:1971–4. https://doi.org/10.1093/rheumatology/keq194

206. Saleem B, Keen H, Goeb V, Parmar R, Nizam S, Hensor EM, *et al.* Patients with RA in remission on TNF blockers: when and in whom can TNF blocker therapy be stopped? *Ann Rheum Dis* 2010;**69**:1636–42. https://doi.org/10.1136/ard.2009.117341

207. Brocq O, Millasseau E, Albert C, Grisot C, Flory P, Roux CH, Euller-Ziegler L. Effect of discontinuing TNFalpha antagonist therapy in patients with remission of rheumatoid arthritis. *Joint Bone Spine* 2009;**76**:350–5. https://doi.org/10.1016/j.jbspin.2008.11.009

208. van der Kooij SM, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Guler-Yuksel M, Zwinderman AH, Kerstens PJ, *et al.* Drug-free remission, functioning and radiographic damage after 4 years of response-driven treatment in patients with recent-onset rheumatoid arthritis. *Ann Rheum Dis* 2009;**68**:914–21. https://doi.org/10.1136/ard.2008.092254

209. Nawata M, Saito K, Nakayamada S, Tanaka Y. Discontinuation of infliximab in rheumatoid arthritis patients in clinical remission. *Mod Rheumatol* 2008;**18**:460–4. https://doi.org/10.1007/s10165-008-0089-1

210. van der Bijl AE, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Ten WS, Han KH, van Krugten MV, *et al.* Infliximab and methotrexate as induction therapy in patients with early rheumatoid arthritis. *Arthritis Rheum* 2007;**56**:2129–34. https://doi.org/10.1002/art.22718

211. Quinn MA, Conaghan PG, O'Connor PJ, Karin Z, Greenstein A, Brown A, *et al.* Very early treatment with infliximab in addition to methotrexate in early, poor-prognosis rheumatoid arthritis reduces magnetic resonance imaging evidence of synovitis and damage, with sustained benefit after infliximab withdrawal. *Arthritis Rheum* 2005;**52**:27–35. https://doi.org/10.1002/art.20712

212. Buch MH, Marzo-Ortega H, Bingham SJ, Emery P. Long-term treatment of rheumatoid arthritis with tumour necrosis factor alpha blockade: outcome of ceasing and restarting biologicals. *Rheumatology (Oxford)* 2004;**43**:243–4. https://doi.org/10.1093/rheumatology/keg454

213. Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, *et al.* Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised Phase III trial. ATTRACT Study Group. *Lancet* 1999;**354**:1932–9. https://doi.org/10.1016/S0140-6736(99)05246-0

214. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013*. Process and methods PMG9. URL: www.nice.org.uk/article/pmg9/resources/non-guidance-guide-to-the-methods-of-technology-appraisal-2013-pdf (accessed 20 October 2015).

215. National Institute for Health Research. *NETSCC, HTA. 23 April 2015. Updated Protocol*. URL: https://njl-admin.nihr.ac.uk/document/download/2007323 (accessed 12 February 2018).

216. Joint Formulary Committee. *British National Formulary* (online) London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com/mc/bnf/current/index.htm (accessed 20 October 2015).

217. UK Clinical Research Network Study Portfolio. *Targeted Ultrasound in Rheumatoid Arthritis (TURA)*. 2016. URL: https://www.ukctg.nihr.ac.uk/trials/trial-details/trial-details?trialId=15763 (accessed 20 October 2015).

218. Dundar Y, Dodd S, Dickson R, Walley T, Haycox A, Williamson PR. Comparison of conference abstracts and presentations with full-text articles in the health technology assessments of rapidly evolving technologies. *Health Technol Assess* 2006;**10**(5). https://doi.org/10.3310/hta10050

219. Bossuyt PM, Lijmer JG, Mol BW. Randomised comparisons of medical tests: sometimes invalid, not always efficient. *Lancet* 2000;**356**:1844–7. https://doi.org/10.1016/S0140-6736(00)03246-3

220. Alfredo Chávez-López M, Naredo E, Carlos Acebes-Cachafeiro J, de Miguel E, Cabero F, Sanchez-Pernaute O, *et al.* [Diagnostic accuracy of physical examination of the knee in rheumatoid arthritis: clinical and ultrasonographic study of joint effusion and BakerOs cyst]. *Rheumatol Clin* 2007;**3**:98–100. https://doi.org/10.1016/S1699-258X(07)73675-6

221. Andersen M, Ellegaard K. Ultrasound colour Doppler is associated with synovial pathology in biopsies from hand joints in rheumatoid arthritis patients: a cross-sectional study. *Ann Rheum Dis* 2014;**73**:678–83. https://doi.org/10.1136/annrheumdis-2012-202669

222. Andonopoulos AP, Yarmenitis S, Sfountouris H, Siamplis D, Zervas C, Bounas A. Baker's cyst in rheumatoid arthritis: an ultrasonographic study with a high resolution technique. *Clin Exp Rheumatol* 1995;**13**:633–6.

223. Baan H, Hoekstra M, Veehof M, Van De Laar M. Ultrasound findings in rheumatoid wrist arthritis highly correlate with function. *Disabil Rehabil* 2011;**33**:729–33. https://doi.org/10.3109/09638288.2010.509459

224. Bajaj S, Lopez-Ben, Lopez-Ben R, Oster R, Alarcon G. Ultrasound detects rapid progression of erosive disease in early rheumatoid arthritis: a prospective longitudinal study. *Skelet Radiol* 2007;**36**:123–8. https://doi.org/10.1007/s00256-006-0196-z

225. Boesen M, Boesen L, Jensen KE, Cimmino MA, Torp-Pedersen S, Terslev L, *et al.* Clinical outcome and imaging changes after intraarticular (IA) application of etanercept or methylprednisolone in rheumatoid arthritis: magnetic resonance imaging and ultrasound-Doppler show no effect of IA injections in the wrist after 4 weeks. *J Rheumatol* 2008;**35**:584–91.

226. Brown AK, Quinn MA, Karim, Quinn MA. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. *Arthritis Rheum* 2006;**54**:3761–73. https://doi.org/10.1002/art.22190

227. Bruyn GA, Naredo E, Möller I, Moragues C, Garrido J, de Bock GH, *et al.* Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis. *Ann Rheum Dis* 2009;**68**:357–61. https://doi.org/10.1136/ard.2008.089243

228. Carotti M, Salaffi F, Manganelli P, Salera D, Simonetti B, Grassi W. Power Doppler sonography in the assessment of synovial tissue of the knee joint in rheumatoid arthritis: a preliminary experience. *Ann Rheum Dis* 2002;**61**:877–82. https://doi.org/10.1136/ard.61.10.877

229. Damjanov N, Radunovic G, Prodanovic S, Vukovic V, Milic V, Simic Pasalic K, *et al.* Construct validity and reliability of ultrasound disease activity score in assessing joint inflammation in RA: comparison with DAS-28. *Rheumatology (Oxford)* 2012;**51**:120–8. https://doi.org/10.1093/rheumatology/ker255

230. da Silva Chakr RM, Brenol JC, Behar M, Mendonça JA, Kohem CL, Monticielo OA, *et al.* Is ultrasound a better target than clinical disease activity scores in rheumatoid arthritis with fibromyalgia? A case-control study. *PLOS ONE* 2015;**10**:e0118620. https://doi.org/10.1371/journal.pone.0118620

231. Dejaco C, Duftner C, Wipfler-Freissmuth E, Weiss H, Graninger WB, Schirmer M. Ultrasound-defined remission and active disease in rheumatoid arthritis: association with clinical and serologic parameters. *Semin Arthritis Rheum* 2012;**41**:761–7. https://doi.org/10.1016/j.semarthrit.2011.09.005

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

232. Di Franco M. Hypovitaminosis D in recent onset rheumatoid arthritis is predictive of reduced response to treatment and increased disease activity: a 12 month follow-up study. *BMC Musculoskelet Disord* 2015;**16**:53. https://doi.org/10.1186/s12891-015-0505-6

233. Dohn UM, Ejbjerg B, Boonen A, Hetland ML, Hansen MS, Knudsen LS, *et al.* No overall progression and occasional repair of erosions despite persistent inflammation in adalimumab-treated rheumatoid arthritis patients: results from a longitudinal comparative MRI, ultrasonography, CT and radiography study. *Ann Rheum Dis* 2011;**70**:252–8. https://doi.org/10.1136/ard.2009.123729

234. Epis O, Filippucci E, Delle Sedie A, De Matthaeis A, Bruschi E. Clinical and ultrasound evaluation of the response to tocilizumab treatment in patients with rheumatoid arthritis: a case series. *Rheumatol Int* 2014;**34**:737–42. https://doi.org/10.1007/s00296-012-2638-3

235. Foltz V, Gandjbakhch F, Etchepare F, Rosenberg C, Tanguy ML, Rozenberg S, *et al.* Power Doppler ultrasound, but not low-field magnetic resonance imaging, predicts relapse and radiographic disease progression in rheumatoid arthritis patients with low levels of disease activity. *Arthritis Rheum* 2012;**64**:67–76. https://doi.org/10.1002/art.33312

236. Freeston JE, Brown AK, Hensor EM, Emery, Freeston JE. Extremity magnetic resonance imaging assessment of synovitis (without contrast) in rheumatoid arthritis may be less accurate than power Doppler ultrasound. *Ann Rheum Dis* 2008;**67**:1351. https://doi.org/10.1136/ard.2007.082743

237. Fukae J, Isobe M, Kitano A, Henmi M, Sakamoto F, Narita A, *et al.* Positive synovial vascularity in patients with low disease activity indicates smouldering inflammation leading to joint damage in rheumatoid arthritis: time-integrated joint inflammation estimated by synovial vascularity in each finger joint. *Rheumatology (Oxford)* 2013;**52**:523–8. https://doi.org/10.1093/rheumatology/kes310

238. Fukae J. Structural deterioration of finger joints with ultrasonographic synovitis in rheumatoid arthritis patients with clinical low disease activity. *Rheumatology (Oxford)* 2014;**53**:1608–12. https://doi.org/10.1093/rheumatology/keu154

239. Funck-Brentano T, Etchepare F, Joulin SJ, Gandjbakch F, Pensec VD, Cyteval C, *et al.* Benefits of ultrasonography in the management of early arthritis: a cross-sectional study of baseline data from the ESPOIR cohort. *Rheumatology* 2009;**48**:1515–19. https://doi.org/10.1093/rheumatology/kep279

240. Funck-Brentano T, Gandjbakhch F, Etchepare F, Jousse-Joulin S, Miquel A, Cyteval C, *et al.* Prediction of radiographic damage in early arthritis by sonographic erosions and power Doppler signal: a longitudinal observational study. *Arthritis Care Res* 2013;**65**:896–902. https://doi.org/10.1002/acr.21912

241. Geng Y, Han J, Deng X, Zhang Z. Presence of power Doppler synovitis in rheumatoid arthritis patients with synthetic and/or biological disease-modifying anti-rheumatic drug-induced clinical remission: experience from a Chinese cohort. *Clin Rheumatol* 2014;**33**:1061–6. https://doi.org/10.1007/s10067-014-2634-y

242. Hameed B, Pilcher J, Heron C, Kiely PD. The relation between composite ultrasound measures and the DAS28 score, its components and acute phase markers in adult RA. *Rheumatology* 2008;**47**:476–80. https://doi.org/10.1093/rheumatology/kem383

243. Harman H, Tekeoglu I, Takçi 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. *Clin Rheumatol* 2015;**34**:1367–74. https://doi.org/10.1007/s10067-015-2926-x

244. Harman H, Tekeoglu I, Kaban N, Harman S. Factors influencing ultrasonographic remission in patients with rheumatoid arthritis. *Rheumatol Int* 2015;**35**:485–91. https://doi.org/10.1007/s00296-014-3177-x

245. Hermann KG, Backhaus M, Schneider U, Labs K, Loreck D, Zuhlsdorf S, *et al.* Rheumatoid arthritis of the shoulder joint: comparison of conventional radiography, ultrasound, and dynamic contrast-enhanced magnetic resonance imaging. *Arthritis Rheum* 2003;**48**:3338–49. https://doi.org/10.1002/art.11349

246. Hmamouchi I, Bahiri R, Srifi N, Aktaou S, Abouqal R, Hajjaj-Hassouni N. A comparison of ultrasound and clinical examination in the detection of flexor tenosynovitis in early arthritis. *BMC Musculoskelet Disord* 2011;**12**:91. https://doi.org/10.1186/1471-2474-12-91

247. Janta I, Naredo E, Martinez-Estupinan L, Nieto JC, De la Torre I, Valor L, *et al.* Patient self-assessment and physician's assessment of rheumatoid arthritis activity: which is more realistic in remission status? A comparison with ultrasonography. *Rheumatology (Oxford)* 2013;**52**:2243–50. https://doi.org/10.1093/rheumatology/ket297

248. Kawashiri SY, Suzuki T, Nakashima Y, Horai Y, Okada A, Iwamoto N, *et al.* Ultrasonographic examination of rheumatoid arthritis patients who are free of physical synovitis: power Doppler subclinical synovitis is associated with bone erosion. *Rheumatology (Oxford)* 2014;**53**:562–9. https://doi.org/10.1093/rheumatology/ket405

249. Kawashiri SY, Suzuki T, Nishino A, Nakashima Y, Horai Y, Iwamoto N, *et al.* Automated Breast Volume Scanner, a new automated ultrasonic device, is useful to examine joint injuries in patients with rheumatoid arthritis. *Mod Rheumatol* 2015;**25**:837–41. https://doi.org/10.3109/14397595.2015.1040226

250. Kelly S. Angiogenic gene expression and vascular density are reflected in ultrasonographic features of synovitis in early Rheumatoid Arthritis: an observational study. *Arthritis Res Ther* 2015;**17**:58. https://doi.org/10.1186/s13075-015-0567-8

251. Klauser A, Demharter J, De Marchi A, Sureda D, Barile A, Masciocchi C, *et al.* Contrast enhanced gray-scale sonography in assessment of joint vascularity in rheumatoid arthritis: results from the IACUS study group. *Eur Radiol* 2005;**15**:2404–10. https://doi.org/10.1007/s00330-005-2884-9

252. Klauser AS, Franz M, Arora R, Feuchtner GM, Gruber J, Schirmer M, *et al.* Detection of vascularity in wrist tenosynovitis: power Doppler ultrasound compared with contrast-enhanced grey-scale ultrasound. *Arthritis Res Ther* 2010;**12**:R209. https://doi.org/10.1186/ar3185

253. Krezja. Ultrasonography of the periarticular changes in patients with early active rheumatoid arthritis. *Med Sci Monit* 1998;**4**:MT366–9.

254. Lillegraven S, Boyesen P, Hammer HB, Ostergaard M, Uhlig T, Sesseng S, *et al.* Tenosynovitis of the extensor carpi ulnaris tendon predicts erosive progression in early rheumatoid arthritis. *Ann Rheum Dis* 2011;**70**:2049–50. https://doi.org/10.1136/ard.2011.151316

255. Lillegraven S, Prince FH, Shadick NA, Bykerk VP, Lu B, Frits ML, *et al.* Remission and radiographic outcome in rheumatoid arthritis: application of the 2011 ACR/EULAR remission criteria in an observational cohort. *Ann Rheum Dis* 2012;**71**:681–6. https://doi.org/10.1136/ard.2011.154625

256. Makinen H, Kautiainen H, Hannonen P, Mottonen T, Leirisalo-Repo M, Laasonen L, *et al.* Sustained remission and reduced radiographic progression with combination disease modifying antirheumatic drugs in early rheumatoid arthritis. *J Rheumatol* 2007;**34**:316–21.

257. Montoro AM, Montoro Alvarez M, Chong OY, Janta I, Gonzalez C, Lopez-Longo J, *et al.* Relation of Doppler ultrasound synovitis versus clinical synovitis with changes in native complement component levels in rheumatoid arthritis patients treated with biologic disease-modifying anti-rheumatic drugs. *Clin Exp Rheumatol* 2015;**33**:141–5.

258. Naredo E, Acebes C, Brito E, de Agustin JJ, de Miguel E, Mayordomo L, *et al.* Three-dimensional volumetric ultrasound: a valid method for blinded assessment of response to therapy in rheumatoid arthritis. *J Rheumatol* 2013;**40**:253–60. https://doi.org/10.3899/jrheum.121103

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

259. Naredo E, Hinojosa M, Valor L, Hernández-Flórez D, Mata-Martínez C, Serrano-Benavente B, *et al.* Does ultrasound-scored synovitis depend on the pharmacokinetics of subcutaneous anti-TNF agents in patients with rheumatoid arthritis? *Rheumatology* 2014;**53**:2088–94. https://doi.org/10.1093/rheumatology/keu248

260. Nordal HH, Brun JG, Halse AK, Jonsson, Brun JG. The neutrophil protein S100A12 is associated with a comprehensive ultrasonographic synovitis score in a longitudinal study of patients with rheumatoid arthritis treated with adalimumab. *BMC Musculoskelet Disord* 2014;**15**:335. https://doi.org/10.1186/1471-2474-15-335

261. Peluso G, Michelutti A, Bosello S, Gremese E, Tolusso B, Ferraccioli G. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. *Ann Rheum Dis* 2011;**70**:172–5. https://doi.org/10.1136/ard.2010.129924

262. Ramirez J, Ruiz-Esquide V, Pomés I, Celis R, Cuervo A, Hernández MV, *et al.* Patients with rheumatoid arthritis in clinical remission and ultrasound-defined active synovitis exhibit higher disease activity and increased serum levels of angiogenic biomarkers. *Arthritis Res Ther* 2014;**16**:R5. https://doi.org/10.1186/ar4431

263. Reiche BE, Ohrndorf S, Feist E, Messerschmidt J, Burmester GR, Backhaus M. Usefulness of power Doppler ultrasound for prediction of re-therapy with rituximab in rheumatoid arthritis: a prospective study of longstanding rheumatoid arthritis patients. *Arthritis Care Res (Hoboken)* 2014;**66**:204–16. https://doi.org/10.1002/acr.22103

264. Schmidt WA, Schicke B, Ostendorf B, Scherer A, Krause A, Walther M. Low-field MRI versus ultrasound: which is more sensitive in detecting inflammation and bone damage in MCP and MTP joints in mild or moderate rheumatoid arthritis? *Clin Exp Rheumatol* 2013;**31**:91–6.

265. Spinella A, Sandri G, Carpenito G, Belletti L, Mascia MT. The discrepancy between clinical and ultrasonographic remission in rheumatoid arthritis is not related to therapy or autoantibody status. *Rheumatol Int* 2012;**32**:3917–21. https://doi.org/10.1007/s00296-011-2259-2

266. Stramare R, Coran A, Faccinetto A, Costantini G, Bernardi L, Botsios C, *et al.* MR and CEUS monitoring of patients with severe rheumatoid arthritis treated with biological agents: a preliminary study. *Radiol Med* 2014;**119**:422–31. https://doi.org/10.1007/s11547-013-0369-5

267. Strunk J, Rumbaur C, Albrecht K, Neumann E, Muller-Ladner U. Linking systemic angiogenic factors (VEGF, angiogenin, TIMP-2) and Doppler ultrasound to anti-inflammatory treatment in rheumatoid arthritis. *Joint Bone Spine* 2013;**80**:270–3. https://doi.org/10.1016/j.jbspin.2012.09.001

268. Szkudlarek M, Court-Payen, Strandberg C, Klarlund M, Klausen T, Ostergaard M, *et al.* Power Doppler ultrasonography for assessment of synovitis in the metacarpophalangeal joints of patients with rheumatoid arthritis: a comparison with dynamic magnetic resonance imaging. *Arthritis Rheum* 2001;**44**:2018–23. https://doi.org/10.1002/1529-0131(200109)44:9<2018::AID-ART350>3.0.CO;2-C

269. Taouli B, Zaim S, Peterfy CG, Lynch JA, Stork A, Guermazi A, *et al.* Rheumatoid arthritis of the hand and wrist: comparison of three imaging techniques. *Am J Roentgenol* 2004;**182**:937–43. https://doi.org/10.2214/ajr.182.4.1820937

270. Terslev L, Ellegaard K, Christensen R, Szkudlarek M, Schmidt WA, Jensen PS, *et al.* Head-to-head comparison of quantitative and semi-quantitative ultrasound scoring systems for rheumatoid arthritis: reliability, agreement and construct validity. *Rheumatology (Oxford)* 2012;**51**:2034–8. https://doi.org/10.1093/rheumatology/kes124

271. Varsamidis K, Varsamidou E, Tjetjis V, Mavropoulos G. Doppler sonography in assessing disease activity in rheumatoid arthritis. *Ultrasound Med Biol* 2005;**31**:739–43. https://doi.org/10.1016/j.ultrasmedbio.2005.02.010

272. Wakefield RJ, O'Connor PJ, Conaghan PG, McGonagle D, Hensor EMA, Gibbon WW, *et al.* Finger tendon disease in untreated early rheumatoid arthritis: a comparison of ultrasound and magnetic resonance imaging. *Arthritis Care Res* 2007;**57**:1158–64. https://doi.org/10.1002/art.23016

273. Watanabe T, Takemura M, Sato M, Sekine A, Fukuoka D, Seishima M, *et al.* Quantitative analysis of vascularization in the finger joints in patients with rheumatoid arthritis using three-dimensional volumetric ultrasonography with power Doppler. *Clin Rheumatol* 2012;**31**:299–307. https://doi.org/10.1007/s10067-011-1811-5

274. Yoshimi R, Ihata A, Kunishita Y, Kishimoto D, Kamiyama R, Minegishi K, *et al.* A novel 8-joint ultrasound score is useful in daily practice for rheumatoid arthritis. *Mod Rheumatol* 2015;**25**:379–85. https://doi.org/10.3109/14397595.2014.974305

275. Zheng G, Wang L, Jia X, Li F, Yan Y, Yu Z, *et al.* Application of high frequency color Doppler ultrasound in the monitoring of rheumatoid arthritis treatment. *Exp Ther Med* 2014;**8**:1807–12. https://doi.org/10.3892/etm.2014.2001

276. Ziswiler HR, Aeberli D. High-resolution ultrasound confirms reduced synovial hyperplasia following rituximab treatment in rheumatoid arthritis. *Rheumatology (Oxford)* 2009;**48**:939–43. https://doi.org/10.1093/rheumatology/kep139

277. Meenagh G, Filippucci E, Delle SA, Riente L, Iagnocco A, Epis O, *et al.* Ultrasound imaging for the rheumatologist. XVIII. Ultrasound measurements. *Clin Exp Rheumatol* 2008;**26**:982–5.

278. Koski JM. Axillar ultrasound of the glenohumeral joint. *J Rheumatol* 1989;**16**:664–7.

279. van Riel PLCM, van Gestel AM, Scott DL, editors. *EULAR Handbook of Clinical Assessments in Rheumatoid Arthritis*. Alphen aan den Rijn: Van Zuiden; 2001.

280. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529–36. https://doi.org/10.7326/0003-4819-155-8-201110180-00009

281. Wakefield RJ, Balint PV, Szkudlarek, Balint PV. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005;**32**:2485–7. [Erratum published in *J Rheumatol* 2006;**33**:440.]

282. Wakefield RJ, Green MJ, Marzo-Ortega H, Conaghan PG, Gibbon WW, McGonagle D, *et al.* Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. *Ann Rheum Dis* 2004;**63**:382–5. https://doi.org/10.1136/ard.2003.007062

283. Newman JS, Laing TJ, McCarthy CJ, Adler RS. Power Doppler sonography of synovitis: assessment of therapeutic response – preliminary observations. *Radiology* 1996;**198**:582–4. https://doi.org/10.1148/radiology.198.2.8596870

284. Schmidt WA, Schmidt H, Schicke B, Gromnica-Ihle E. Standard reference values for musculoskeletal ultrasonography. *Ann Rheum Dis* 2004;**63**:988–94. https://doi.org/10.1136/ard.2003.015081

285. Nakagomi D, Ikeda K, Okubo A, Iwamoto T, Sanayama Y, Takahashi K, *et al.* Ultrasound can improve the accuracy of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis to predict the requirement for methotrexate treatment. *Arthritis Rheum* 2013;**65**:890–8. https://doi.org/10.1002/art.37848

286. Lund PJ, Heikal A, Maricic MJ, Krupinski EA, Williams CS. Ultrasonographic imaging of the hand and wrist in rheumatoid arthritis. *Skeletal Radiol* 1995;**24**:591–6. https://doi.org/10.1007/BF00204858

287. Rutjes AW, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007;**11**(50). https://doi.org/10.3310/hta11500

288. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 1 Survey

There were few publications regarding US and its use in treatment decisions (see *Chapter 1*). A survey was conducted of UK rheumatology units to investigate how US was being used in practice to influence therapy. The BSR publicised the survey to UK rheumatology units. Additionally, the BSR requested that respondents be questioned about the use of US in the diagnosis of RA.

The survey was available to clinicians from December 2015 to February 2016.

Only 31 responses were received by the end of February 2016. Survey respondents were self-selecting and so the sample may have been biased.

Survey form

Do you use US for diagnosis of RA, yes/no?

Do you use US for monitoring synovitis in RA, yes/no?

If you answered yes to the previous question, how frequently are patients monitored?

and what information is routinely collected (e.g. number of joints, is a particular joint count used)?

Do you use US to make decisions regarding RA therapy, yes/no?

If so, is this for decisions to start/stop medication?

Is it used for dose adjustments (e.g. tapering medication)?

Is it used for decisions regarding cDMARDs, biologics, and/or steroids?

Who conducts US, rheumatologists or radiologists or other allied health professionals (e.g. physiotherapists, nurses, podiatrists)?

Have the people conducting US received formal training in US to detect synovitis?

In which Rheumatology Unit are you based?

This survey is to provide background information for a project funded by the NIHR Health Technology Assessment programme. Survey results will be aggregated for publication and no identifying information of participants will be published. The project will be published in the *Health Technology Assessment* journal series. Visit the HTA programme website for more details www.nets.nihr.ac.uk/programmes/hta

Any views and opinions expressed in *Health Technology Assessment* journal articles are those of the authors and do not necessarily reflect those of the Department of Health.

Survey answers

Ultrasound was used in the diagnosis of RA by 27 out of 31 (87%) respondents.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Twenty respondents said they used US for monitoring synovitis (*Table 15*). Additionally, one respondent said that their unit planned to use US in the future. Of the 20 respondents using US for monitoring synovitis, five said that monitoring was routine, with two stating that monitoring was routine only in early arthritis. When monitoring was routine, this was every 3 or 6 months or annually. US assessment was stated explicitly to be used for cases of uncertainty regarding synovitis (12/20) or symptoms change (1/20), for distinguishing synovitis from other pathology (1/20) or for making a treatment decision or monitoring treatment (7/20). Information collected was stated by four respondents to be specific to the clinical or treatment question for which the patient was referred for US. Three respondents stated that information was collected on the presence or absence of synovitis and four stated that the number of active joints was recorded. Most respondents did not use a particular joint count, but there was one mention each of a seven-joint score, the DAS28 and 34 joints and three respondents assessed wrist and hand joints. Four respondents mentioned that grading of GSUS and PDUS was used. US was reported as being conducted by rheumatologists (23/31), radiologists (16/31) and allied professionals (6/31).

Twenty-seven respondents stated that US was used for making treatment decisions, with 20 out of 27 respondents using US for making decisions about starting or stopping medication and 4 out of 27 respondents using US for making decisions about starting medication (not stopping). US was used in making decisions around dose adjustment by 19 of the 27 respondents, with an additional two respondents stating that US was used to make decisions around escalating (but not tapering) medication. A further respondent stated that US may be used in the future for making treatment decisions.

Ultrasound was used for making treatment decisions around cDMARDs, bDMARDs and steroids by 17 respondents, for bDMARDs by three respondents and for cDMARDs and steroids by three respondents. Two respondents reported the use of US when treatment decisions were difficult for selected patients. Two respondents commented that there were plans to use US in the future for making decisions about treatment with bDMARDs.

From the responses, it can be concluded that some units do use US for making treatment decisions, but the small sample size and potential for bias mean that the results cannot be generalised across all UK rheumatology units.

TABLE 15 Survey results

Question	Yes, *n*/*N* (%)	No, *n*/*N* (%)	Don't know, *n*/*N* (%)
Do you use US for monitoring synovitis in RA?	20/31 (65)	11/31 (35)	
Do you use US to make decisions regarding RA therapy?	27/31 (87)	4/31(13)	
Have the people conducting US received formal training in US to detect synovitis?	25/31 (81)	1/31 (3)	5/31 (16)

Appendix 2 Patient involvement

The National Rheumatoid Arthritis Society were contacted about patient involvement in the study. It provided the names of two patients with RA who were willing to be contacted about being involved. Both patients were sent a draft of the plain English summary and were asked to provide feedback. The full report was available to provide further information. The two patients were also invited to provide comments regarding the patient experience of US and suggestions for future research priorities.

One patient did not respond. The other patient provided feedback on the plain English summary and contributed text on the patient experience of US. This feedback was helpful and was deemed to improve the accessibility of the plain English summary to a lay audience. The text of the plain English summary was amended in accordance with the patient's suggestions. The background section of the report was amended based on information provided by the patient advisor on the patient experience of US.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 3 Literature search strategies

There were three phases of searches: (1) initial searches to identify key words, (2) the phase 1 scoping searches and (3) the phase 2 comprehensive searches. The scoping searches were conducted to determine whether or not evidence was available on US for monitoring synovitis in RA and, therefore, whether or not phase 2 of the project could be justified and whether or not any cost-effectiveness analyses could be conducted.

The phase 1 searches identified diagnostic accuracy and prognostic studies relevant to the monitoring of synovitis by US and so comprehensive searches for the diagnostic, prognostic and treatment effectiveness section of the project were conducted as planned. However, the phase 1 scoping searches revealed a lack of data to populate the planned cost-effectiveness model. Following advice from the clinical advisors and consultation with the NIHR HTA programme, it was decided that an analysis of the costs and benefits associated with US would be of value to the clinical community. However, attempting to achieve this using the model that informed the recent NICE multiple technology appraisal[40] was not deemed sensible as this contained many elements that would be unnecessary when focusing on current decision problem and would add uncertainty to the results. Therefore, a simpler model was constructed, which was subject to sensitivity analyses and threshold analyses.

Initial searches to identify key words

Following the peer reviewer comments on the protocol,[215] initial scoping searches for reviews and diagnostic accuracy studies were carried out on 12 March 2015 to predict the size of the evidence base. Searches for existing guidelines (national and international) were considered in the scoping searches. Existing systematic reviews were searched for in a selected number of databases by applying a specific reviews search filter. Focused diagnostic accuracy studies were searched for by applying a specific diagnostic filter. This identified 114 records, indicating that a broader search would be useful.

Phase 1 searches

Based on the initial scoping searches, design filters were not combined with the search strategies. Date and English-language limits were not applied.

The following databases were searched on 12 March 2015:

- MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid)
- EMBASE (via Ovid)
- The Cochrane Library (via Wiley Online Library)
- Science Citation Index Expanded (via Web of Science)
- Science Citation Index and Conference Proceedings Index (via Web of Science).

The number of records from each database (without date limits or application of a study design filter) is shown in *Table 16*.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 16 Results of the database searches, 12 March 2015

Approach	Source	Number of records
Electronic database	MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations	459
	EMBASE	1313
	HTA database	0
	DARE	3
	CDSR	0
	CENTRAL	35
	NHS EED	0
	Science Citation Index Expanded and Science Citation Index and Conference Proceedings Index	734
Electronic database and trials registry	ClinicalTrials.gov	123
Total	Retrieved	2677
	Unique	1742
Conference abstracts via Web of Science	European League Against Rheumatism Abstract Archive	18
	American College of Rheumatology and Association of Rheumatology Health Professionals	52
	OMERACT conference proceedings	6
Total	Retrieved	76
	Unique	25
Overall total	Unique records	1767

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid) (1946–2015)

Date searched: 12 March 2015.

Search strategy

1. exp Arthritis, Rheumatoid/
2. rheumatoid arthritis.tw.
3. or/1-2
4. exp Synovitis/
5. synovitis.tw.
6. ((synovial or synovium) adj5 inflam$).tw.
7. or/4-6
8. exp Ultrasonography/
9. ultrasound.tw.
10. ultrason$.tw.
11. sonography.tw.
12. echography.tw.
13. or/8-12
14. 3 and 7 and 13

EMBASE (via Ovid) (1974 to 11 March 2015)

Date searched: 12 March 2015.

Search strategy

1. exp rheumatoid arthritis/
2. rheumatoid arthritis.tw.
3. or/1-2
4. exp synovitis/
5. synovitis.tw.
6. ((synovial or synovium) adj5 inflam$).tw.
7. or/4-6
8. exp echography/
9. ultrasound.tw.
10. ultrason$.tw.
11. sonography.tw.
12. echography.tw.
13. or/8-12
14. 3 and 7 and 13

The Cochrane Library: Cochrane Database of Systematic Reviews (1996–2015), Cochrane Central Register of Controlled Trials (1898–2015), Health Technology Assessment database (1989–2015), Database of Abstracts of Reviews of Effects (1946–2014) and NHS Economic Evaluation Database (1968–2014) (via Wiley Online Library)

Date searched: 12 March 2015.

Search strategy

#1 MeSH descriptor: [Arthritis, Rheumatoid] explode all trees

#2 rheumatoid arthritis:ti,ab,kw

#3 #1 or #2

#4 MeSH descriptor: [Synovitis] explode all trees

#5 synovitis:ti,ab,kw

#6 ((synovial or synovium) next/5 inflam*):ti,ab,kw

#7 #4 or #5 or #6

#8 ultrasound:ti,ab,kw

#9 ultrason*:ti,ab,kw

#10 sonography:ti,ab,kw

#11 or #8-#10

#12 #3 and #7 and #11

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Science Citation Index Expanded and Science Citation Index and Conference Proceedings Index (1990–2015) (via Web of Science)

Date searched: 12 March 2015.

Search strategy

#10 #9 AND #4 AND #1

#9 #8 OR #7 OR #6 OR #5

#8 TOPIC: (echography)

#7 TOPIC: (sonography)

#6 TOPIC: (ultrason*)

#5 TOPIC: (ultrasound)

#4 #3 OR #2

#3 TOPIC: (((synovial or synovium) NEAR/5 inflam*))

#2 TOPIC: (synovitis)

#1 TOPIC: (rheumatoid arthritis)

ClinicalTrials.gov (US National Institutes of Health)

Date searched: 12 March 2015.

123 studies found for ultrasound | arthritis.

122 studies found for ultrasonography | arthritis.

122 studies found for sonography | arthritis.

122 studies found for echography | arthritis.

European League Against Rheumatism Abstract Archive (via Web of Science)

URL: http://scientific.sparx-ip.net/archiveeular/ (accessed 20 March 2015); published in *Annals of the Rheumatic Diseases*.

Date searched: 20 March 2015.

Number of results: 18.

Search strategy

#4 #3 AND #2 AND #1

#3 TOPIC: (ultrasound) OR TOPIC: (ultrason*) OR TOPIC: (sonography) OR TOPIC: (echography)

#2 TOPIC: (((synovial or synovium) NEAR/5 inflam*)) OR TOPIC: (synovitis)

#1 PUBLICATION NAME: (ann rheum dis)

American College of Rheumatology and Association of Rheumatology Health Professionals (Web of Science)

URL: http://acrabstracts.org/ (accessed 8 February 2018); published in *Arthritis & Rheumatology*.

Date searched: 20 March 2015.

Number of results: 52.

Search strategy

#4 #3 AND #2 AND #1

#3 TOPIC: (ultrasound) OR TOPIC: (ultrason*) OR TOPIC: (sonography) OR TOPIC: (echography)

#2 TOPIC: ((((synovial or synovium) NEAR/5 inflam*))) OR TOPIC: ((synovitis))

#1 SO=(ARTHRITIS "AND" RHEUMATISM)

An update search was carried out in May 2015, with 32 unique records retrieved (*Table 17*). Added to the 1767 records identified from the searches conducted in March 2015, this gave a total of 1799 records from the phase 1 electronic database searches.

The study selection criteria were refined based on the phase 1 search results. The protocol[215] had stated that 'For studies of diagnostic accuracy, study designs will be accepted into the review according to the hierarchy of evidence published by Merlin *et al.*[159]' As few studies were identified with diagnostic accuracy data, it was decided that diagnostic studies providing sensitivity or specificity data would be included, even if this was not the highest level of evidence according to the hierarchy,[159] that is, diagnostic test accuracy studies with an independent, blinded comparator of a valid reference standard, tested on consecutive patients.

TABLE 17 Results of the update database searches, 13 May 2015

Approach	Source	Number of records
Electronic database	MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations	11
	EMBASE	20
	HTA database	0
	DARE	0
	CDSR	0
	CENTRAL	0
	NHS EED	0
	Science Citation Index Expanded and Science Citation Index and Conference Proceedings Index	21
	PubMed	24
Conference abstracts via Web of Science	European League Against Rheumatism Abstract Archive	1
	American College of Rheumatology and Association of Rheumatology Health Professionals	0
	OMERACT conference abstracts proceedings	3
Total	Retrieved	80
Final total	Records from all databases with duplicates removed	32 unique to add to database

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The phase 1 searches did not identify many prognostic studies or studies about response to treatment. Studies of any level according to the hierarchy of prognostic studies[159] were included as well as studies with any outcome relating the use of US to treatment decisions or incorporating US to predict treatment adherence, response, failure, relapse on, or following, discontinuation of treatment and surveys looking at how US is used in treatment decisions in practice.

Phase 2 searches

The phase 1 scoping searches had identified that relevant diagnostic and prognostic studies were available and therefore the phase 2 searches were carried out. However, the phase 1 searches also revealed a lack of data to populate the planned cost-effectiveness model. As such, it was proposed that a simpler model would be constructed.

As the phase 1 searches had been successful in identifying relevant studies and had not incorporated design filters or date and English-language limits, it was decided that the same search strategies would be used, slightly broadened, for the clinical effectiveness searches in phase 2. Additionally, a cost-effectiveness filter was applied for the cost-effectiveness searches.

The following databases were searched for clinical effectiveness, cost-effectiveness and dose modification studies between 22 October and 6 November 2015:

- BIOSIS Previews
- MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid)
- EMBASE (via Ovid)
- The Cochrane Library (via Wiley Online Library)
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (via EBSCO*host*)
- PsycINFO (via Ovid)
- Science Citation Index Expanded (via Web of Science)
- Science Citation Index and Conference Proceedings Index (Web of Science)
- TOXLINE (via ProQuest)
- ProQuest Dissertations & Theses A&I (formerly Dissertation Abstracts International)
- ProQuest Dissertations & Theses – UK & Ireland
- EconLit.

The following trial registries, conference proceedings and websites were searched between 26 and 27 October 2015:

- American College of Rheumatology and Association of Rheumatology Health Professionals
- European League Against Rheumatism Abstract Archive
- ClinicalTrials.gov
- Current Controlled Trials
- NICE Evidence Search
- BSR
- Arthritis Research UK
- British Pain Society
- Arthritis and Musculoskeletal Alliance
- National Guideline Clearinghouse
- Royal College of Physicians
- Royal College of Radiologists
- Royal College of Pathologists
- Royal College of Surgeons
- NRAS.

The phase 2 clinical effectiveness searches identified an additional 925 records. Added to the results of the phase 1 searches this gave a total of 2724 unique records from the electronic database and supplementary searches (study selection for the clinical effectiveness records is shown in *Figure 1*). In addition, the cost-effectiveness searches identified 226 records and a parameter search for dose modification, which was not restricted to imaging studies, identified 54 records (study selection for the cost-effectiveness review is shown in *Chapter 4*).

TABLE 18 Results of the database searches, October–November 2015

Approach	Source	Number of records
Electronic database	MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations	596
	EMBASE	1574
	CINAHL	132
	DARE	2
	CDSR	0
	CENTRAL	1
	NHS EED	0
	EconLit	0
	PsycINFO	1
	TOXLINE	2
	BIOSIS Previews	508
	Science Citation Index Expanded and Science Citation Index and Conference Proceedings Index	975
	ProQuest Dissertations & Theses A&I	434
	ProQuest Dissertations & Theses – UK & Ireland	23
Conference abstracts via Web of Science	European League Against Rheumatism Abstract Archive	0
	American College of Rheumatology and Association of Rheumatology Health Professionals	3
	OMERACT conference abstracts	0
Grey literature	Current Controlled Trials	1
	ClinicalTrials.gov	23
	BSR	0
	Royal College of Physicians	0
	Royal College of Radiologists	3
	Royal College of Surgeons	0
	NRAS	41
	Royal College of Pathologists	0
	Arthritis Research UK	16
	British Pain Society	0
	Arthritis and Musculoskeletal Alliance	1
	National Guideline Clearinghouse	19
	NICE Evidence Search	33
Total	Retrieved total	4388
	Unique total	1205

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Search strategies

BIOSIS Previews (via Web of Science) (1969–2015)

Date searched: 6 November 2015.

Search strategy

#1 TOPIC: (rheumat* NEAR/5 (nodule or arthritis))

#2 TOPIC: =(felty* NEAR/2 syndrome)

#3 TOPIC: =(Caplan* NEAR/2 syndrome)

#4 TOPIC: (Sjogren* NEAR/2 syndrome)

#5 TOPIC: (Sicca NEAR/2 syndrome)

#6 TOPIC: Still* disease

#7 TOPIC: Bechterew* disease

#8 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7

#9 TOPIC: Synovitis

#10 TOPIC: ((Synovial or synovium) NEAR/5 (inflam* or hypertrophy))

#11 #9 OR #10

#12 TOPIC: Ultrasound

#13 TOPIC: Ultrason*

#14 TOPIC: Sonography

#15 TOPIC: Echography

#16 TOPIC: Ultrasonic

#17 #12 OR #13 OR #14 OR #15 OR #16

#18 #8 AND #11 #17

Cumulative Index to Nursing and Allied Health Literature (via EBSCO*host*) (1969–2015)

Date searched: 22 October 2015.

Search strategy

S1 (MH "Arthritis, Rheumatoid+")

S2 TX (Rheumat* N5 (nodule or arthritis))

S3 (Felty* N2 syndrome)

S4 TX (Sjogren* N2 syndrome)

S5 TX (Sicca N2 syndrome)

S6 TX Still* disease

S7 TX Bechterew*

S8 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7

S9 (MH "Synovitis")

S10 TS=((Synovial or synovium) NEAR/5 (inflam* or hypertrophy))

S11 #9 OR #10

S12 TX Ultrasound

S13 TX Ultrason*

S14 TX Sonography

S15 TX Echography

S16 TX Ultrasonic

S17 S12 OR S13 OR S14 OR S15 OR S16

S18 S8 AND S11 AND S17

The Cochrane Library: Cochrane Database of Systematic Reviews (1996–2015), Cochrane Central Register of Controlled Trials (1898–2015), Health Technology Assessment database (1989–2015), Database of Abstracts of Reviews of Effects (1946–2014) and NHS Economic Evaluation Database (1968–2014) (via Wiley Online Library)

Date searched: 22 October 2015.

Search strategy

#1 MeSH descriptor: [Arthritis, Rheumatoid] explode all trees

#2 (Rheumat* next/5 (nodule or arthritis)):ti,ab,kw

#3 (felty* next/2 syndrome):ti,ab,kw

#4 (caplan* next/2 syndrome):ti,ab,kw

#5 (Sjogren* next/2 syndrome):ti,ab,kw

#6 (Sicca next/2 syndrome):ti,ab,kw

#7 Still* disease:ti,ab,kw

#8 Bechterew* disease:ti,ab,kw

#9 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

#10 MeSH descriptor: [Synovitis] explode all trees

#11 synovitis:ti,ab,kw

#12 synovitis:ti,ab,kw

#13 #10 OR #11 OR #12

#14 MeSH descriptor: [Ultrasonography] explode all trees

#15 Ultrasound:ti,ab,kw

#16 Ultrason*:ti,ab,kw

#17 Sonography:ti,ab,kw

#18 Echography:ti,ab,kw

#19 Ultrasonic:ti,ab,kw

#20 #14 or #15 or #16 or #17 or #18 or #19

#21 #9 and #13 and #20

ClinicalTrials.gov (US National Institutes of Health)

Date searched: 26 October 2015.

91 studies found for ultrasound l rheumatoid arthritis.

Checked for duplicates: 23 imported.

Current Controlled Trials

Date searched: 27 October 2015.

Search strategy

1. text search: Synovitis condition: Rheumatoid Arthritis AND Interventions: ultrasonography 0 results found
2. Condition: Rheumatoid Arthritis AND Interventions: ultrasonography 1 result found
3. Condition: Rheumatoid Arthritis AND Interventions: ultrasonound 0 results found
4. Condition: Rheumatoid Arthritis AND Interventions: sonography 0 results found
5. Condition: Rheumatoid Arthritis AND Interventions: Echography 0 results found

EMBASE (via Ovid) (1974 to 20 October 2015)

Date searched: 20 October 2015.

Search strategy

1. exp rheumatoid arthritis/
2. (rheumat$ adj5 (nodule or arthritis)).tw.
3. felty$ adj2 syndrome).tw
4. (caplan$ adj2 syndrome).tw
5. (rheumat$ adj2 (nodule or arthritis)).tw

6. (sjogren$ adj2 syndrome).tw
7. (sicca adj2 syndrome).tw
8. still$ disease.tw
9. bechterew$ disease.tw
10. or/1-9
11. exp synovitis/
12. synovitis.tw.
13. ((Synovial or synovium) adj5 (inflam$ or hypertrophy)).tw
14. or/11-13
15. exp echography/
16. ultrasound.tw.
17. ultrason$.tw.
18. sonography.tw.
19. echography.tw.
20. Ultrasonic.tw
21. or/15-20
22. 11 and 14 and 21

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid) (1946–2015)

Date searched: 21 October 2015.

Search strategy

1. exp Arthritis, Rheumatoid/
2. (Rheumat$ adj5 (nodule or arthritis)).tw
3. (Felty$ adj2 syndrome).tw
4. (Caplan$ adj2 syndrome).tw
5. (Sjogren$ adj2 syndrome).tw
6. (Sicca adj2 syndrome).tw
7. Still$ disease.tw
8. Bechterew$ disease.tw
9. or/1-8
10. Synovitis/
11. Synovitis.tw.
12. ((Synovial or synovium) adj5 (inflam$ or hypertrophy)).tw.
13. or/10-12
14. exp Ultrasonography/
15. Ultrasound.tw
16. Ultrason$.tw.
17. Sonography.tw
18. Echography.tw.
19. Ultrasonic.tw
20. or/14-19
21. 9 and 13 and 20

NICE Evidence Search

URL: www.evidence.nhs.uk/ (accessed 8 February 2018).

Date searched: 26 October 2015.

52 studies found for "rheumatoid arthritis" l "Synovitis" l "ultrasonography"

Checked for duplicates: 33 imported

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Science Citation Index Expanded and Science Citation Index and Conference Proceedings Index (via Web of Science) (1900–2015)

Date searched: 22 October 2015.

Search strategy

#18 #8 AND #11 AND #17

#17 #12 OR #13 OR #14 OR #15 OR #16

#16 TOPIC: Ultrasonic

#15 TOPIC: Echography

#14 TOPIC: Sonography

#13 TOPIC: Ultrason*

#12 TOPIC: Ultrasound

#11 #9 OR #10

#10 TOPIC: ((Synovial or synovium) NEAR/5 (inflam* or hypertrophy))

#9 TOPIC: Synovitis

#8 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7

#7 TOPIC: Bechterew* disease

#6 TOPIC: Still* disease

#5 TOPIC: (Sicca NEAR/2 syndrome)

#4 TOPIC: (Sjogren* NEAR/2 syndrome)

#3 TOPIC: (Caplan* NEAR/2 syndrome)

#2 TOPIC: (Felty* NEAR/2 syndrome)

#1 TOPIC: (rheumat* NEAR/5 (nodule or arthritis))

TOXLINE (via ProQuest) (1999–2015)

Date searched: 6 November 2015.

Search strategy

S17 S7 AND S10 AND S16

S16 S11 OR S12 OR S13 OR S14 OR S15

S15 Ultrasonic

S14 Echography

S13 Sonography

S12 Ultrason*

S11 Ultrasound

S10 S8 OR S9

S9 ((Synovial or synovium) NEAR/5 (inflam* or hypertrophy))

S8 Synovitis

S7 S1 OR S2 OR S3 OR S4 OR S5 OR S6

S6 Still* disease

S5 (Sicca NEAR/2 syndrome)

S4 (Caplan NEAR/2 syndrome)

S3 (Felty* NEAR/2 syndrome)

S2 (Rheumat* NEAR/5 (nodule or arthritis))

S1 Arthritis, Rheumatoid

PsycINFO (via Ovid) (1806 to October week 3 2015)

Date searched: 21 October 2015.

Search strategy

1. exp Rheumatoid Arthritis/
2. (Rheumat$ adj5 (nodule or arthritis)).tw.
3. (Felty$ adj2 syndrome).tw.
4. (Caplan$ adj2 syndrome).tw.
5. (Sjogren$ adj2 syndrome).tw.
6. (Sicca adj2 syndrome).tw.
7. Still$ disease.tw.
8. Bechterew$ disease.tw.
9. or/1-8
10. Synovitis.tw
11. ((Synovial or synovium) adj5 (inflam$ or hypertrophy)).tw
12. or/10-11
13. exp Ultrasound/
14. Ultrasound.tw
15. Ultrason$.tw
16. Sonography.tw
17. Echography.tw
18. Ultrasonic.tw
19. or/13-18
20. 9 and 12 and 19

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

ProQuest Dissertations & Theses A&I (via ProQuest) (1743–2015)
Date searched: 26 October 2015.

Search strategy
S1 (Rheumat* NEAR/5 (nodule or arthritis))

S2 (Felty* NEAR/2 syndrome)

S3 (Caplan* NEAR/2 syndrome)

S4 (Sjogren* NEAR/2 syndrome)

S5"Still* disease"

S6 "Bechterew* disease"

S7 S1 OR S2 OR S3 OR S4 OR S5 OR S6

S8 Synovitis

S9 ((Synovial or synovium) NEAR/5 (inflam* or hypertrophy))

S10 s8 or s9

S11 Ultrasound

S12 Ultrason*

S13 Sonography

S14 Echography

S15 Ultrasonic

S16 S11 OR S12 OR S13 OR S14 OR S15

S17 S7 AND S10 AND S16

ProQuest Dissertations & Theses – UK & Ireland (via ProQuest) (1986–2015)
Date searched: 26 October 2015.

Search strategy
S1 (Rheumat* NEAR/5 (nodule or arthritis))

S2 (Felty* NEAR/2 syndrome)

S3 (Caplan* NEAR/2 syndrome)

S4 (Sjogren* NEAR/2 syndrome)

S5 "Still* disease"

S6 "Bechterew* disease"

S7 S1 OR S2 OR S3 OR S4 OR S5 OR S6

S8 Synovitis

S9 ((Synovial or synovium) NEAR/5 (inflam* or hypertrophy))

S10 S8 or S9

S11 Ultrasound

S12 Ultrason*

S13 Sonography

S14 Echography

S15 Ultrasonic

S16 S11 OR S12 OR S13 OR S14 OR S15

S17 S7 AND S10 AND S16

European League Against Rheumatism Abstract Archive (via Web of Science)

URL: http://scientific.sparx-ip.net/archiveeular/ (accessed 9 November 2015); published in *Annals of the Rheumatic Diseases*.

Date searched: 9 November 2015.

Search strategy

#4 #3 AND #2 AND #1

#3 TOPIC: (ultrasound) OR TOPIC: (ultrason*) OR TOPIC: (sonography) OR TOPIC: (echography)

#2 TOPIC: (((synovial or synovium) NEAR/5 inflam*)) OR TOPIC: (synovitis)

#1 PUBLICATION NAME: (ann rheum dis)

American College of Rheumatology and Association of Rheumatology Health Professionals (via Web of Science)

URL: http://acrabstracts.org/ (accessed 8 February 2018); published in *Arthritis & Rheumatology*.

Date searched: 9 November 2015.

Search strategy

#4 #3 AND #2 AND #1

#3 TOPIC: (ultrasound) OR TOPIC: (ultrason*) OR TOPIC: (sonography) OR TOPIC: (echography)

#2 TOPIC: ((((synovial or synovium) NEAR/5 inflam*))) OR TOPIC: ((synovitis))

#1 SO=(ARTHRITIS "AND" RHEUMATISM)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

British Society for Rheumatology
URL: www.rheumatology.org.uk/

Date searched: 27 October 2015.

0 studies found for Ultrasound.

0 studies found for Ultrasonography.

0 studies found for Sonography.

0 studies found for Echography.

0 studies found for Synovitis.

Royal College of Physicians
URL: www.rcplondon.ac.uk/

Date searched: 27 October 2015.

0 studies found for Ultrasound synovitis.

0 studies found for Ultrasonography.

0 studies found for Sonography.

0 studies found for Echography.

0 studies found for Synovitis.

Royal College of Radiologists
URL: www.rcr.ac.uk/

Date searched: 27 October 2015.

0 studies found for Rheumatoid arthritis ultrasound.

0 studies found for Rheumatoid arthritis ultrasonography.

0 studies found for Rheumatoid arthritis sonography.

0 studies found for Rheumatoid arthritis echography.

0 studies found for Rheumatoid arthritis synovitis.

Royal College of Surgeons
URL: www.rcseng.ac.uk

Date searched: 27 October 2015.

0 results found for Synovitis.

Royal College of Pathologists
URL: www.rcpath.org/

Date searched: 27 October 2015.

0 studies found for Ultrasound synovitis

0 studies found for Ultrasonography synovitis

0 studies found for Sonography synovitis

0 studies found for Echography synovitis

National Rheumatoid Arthritis Society
URL: www.nras.org.uk (accessed 18 March 2015).

Date searched: 27 October 2015.

25 results found for Ultrasound synovitis.

6 studies found for Ultrasonography synovitis.

5 studies found for Sonography synovitis.

5 studies found for Echography synovitis.

Arthritis Research UK
URL: www.arthritisresearchuk.org/

Date searched: 27 October 2015.

13 studies found for Ultrasound synovitis (filter by research).

1 study found for Ultrasonography synovitis.

1 study found for Sonography synovitis.

1 study found for Echography synovitis.

British Pain Society
URL: www.britishpainsociety.org/

Date searched: 27 October 2015.

0 studies found for Rheumatoid arthritis

National Guideline Clearinghouse
URL: www.guideline.gov/

Date searched: 27 October 2015.

19 studies found for "Rheumatoid arthritis ultrasonography".

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Cost-effectiveness searches

Cumulative Index to Nursing and Allied Health Literature (via EBSCO*host*) (1969–2015)

Date searched: 22 October 2015.

Search strategy

S1 (MH "Economics+")

S2 (MH "Financial Management+")

S3 (MH "Financial Support+")

S4 (MH "Financing, Organized+")

S5 (MH "Business+")

S6 S2 OR S3 OR S4 OR S5

S7 S1 NOT S6

S8 MH Health resource allocation

S9 MH Health resource utilization

S10 S8 OR S9

S11 S7 OR S10

S12 TX (Cost or costs or economic* or pharmacoeconomic* or price* or pricing*)

S13 S11 OR S12

S14 PT Editorial

S15 PT Letter

S16 PT News

S17 S14 OR S15 OR S16

S18 S13 NOT S17

S19 (MH "Animal Studies")

S20 S18 NOT S19

S21 (MH "Arthritis, Rheumatoid+")

S22 TX (Rheumat* N5 (nodule or arthritis))

S23 (Felty* N2 syndrome)

S24 TX (Caplan* N2 syndrome)

S25 TX (Sjogren* N2 syndrome)

S26 TX (Sicca N2 syndrome)

S27 TX Still* disease

S28 TX Bechterew* disease

S29 S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28

S30 (MH "Ultrasonography+")

S31 TX Ultrasound

S32 TX Ultrason*

S33 TX Sonography

S34 TX Echography

S35 TX Ultrasonic

S36 S30 OR S31 OR S32 OR S33 OR S34 OR S35

S37 S18 AND S29 AND S36

EMBASE (via Ovid) (1974 to 20 October 2015)

Date searched: 21 October 2015.

Search strategy

1. Socioeconomics/
2. Cost benefit analysis/
3. Cost-effectiveness analysis/
4. Cost of illness/
5. Cost control/
6. Economic aspect/
7. Financial management/
8. Health care cost/
9. Health care financing/
10. Health economics/
11. Hospital cost/
12. (fiscal or financial or finance or funding).tw
13. Cost minimisation analysis/
14. (cost adj estimate$).mp
15. (cost adj variable$).mp
16. (unit adj cost$).mp
17. or/1–16
18. exp rheumatoid arthritis/
19. (rheumat$adj5 (nodule or arthritis)).tw synovitis.tw.
20. (felty$adj2 syndrome).tw
21. (Caplan$adj2 syndrome).tw
22. (Rheumat$adj2 (nodule or arthritis)).tw

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

23. (Sjogren$adj2 syndrome).tw
24. (Sicca adj2 syndrome).tw
25. Still$disease.tw
26. Bechterew$disease.tW
27. or/18–26
28. exp echography/
29. ultrasound.tw.
30. ultrason$.tw.
31. sonography.tw.
32. echography.tw.
33. Ultrasonic.tw
34. or/28–33
35. 17 and 28 and 34

MEDLINE, MEDLINE Daily and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid) (1946–2015)

Date searched: 21 October 2015.

Search strategy

1. Economics/
2. "costs and cost analysis"/
3. Cost-effectiveness analysis/
4. Cost–benefit analysis/
5. Cost control/
6. Cost savings/
7. Cost of illness/
8. Cost sharing/
9. "deductibles and coinsurance"/
10. Medical savings accounts/
11. Health care costs/
12. Direct service costs/
13. Drug costs/
14. Employer health costs/
15. Hospital costs/
16. Health expenditures/
17. Capital expenditures/
18. Value of life/
19. exp economics, hospital/
20. exp economics, medical/
21. Economics, nursing/
22. Economics, pharmaceutical/
23. exp "fees and charges"/
24. exp budgets/
25. (Low adj cost).mp.
26. (High adj cost).mp.
27. (Health?care adj cost$).mp.
28. (Fiscal or funding or financial or finance).tw.
29. (Cost adj estimate$).mp.
30. (Cost adj variable).mp.
31. (Unit adj cost$).mp.
32. (Economic$or pharmacoeconomic$or price$or pricing).tw.
33. or/1–32

34. exp Arthritis, Rheumatoid/
35. (Rheumat$adj5 (nodule or arthritis)).tw
36. (Felty$adj2 syndrome).tw.
37. (Caplan$adj2 syndrome).tw.
38. (Sjogren$adj2 syndrome).tw.
39. (Sicca adj2 syndrome).tw.
40. Still$disease.tw.
41. Bechterew$disease.tw.
42. or/34–41
43. exp Ultrasonography/
44. Ultrasound.tw.
45. Ultrason$.tw.
46. Sonography.tw.
47. Echography.tw.
48. Ultrasonic.tw.
49. or/43–48
50. 33 and 42 and 49

Science Citation Index Expanded and Science Citation Index and Conference Proceedings Index (via Web of Science) (1900–2015)

Date searched: 23 October 2015.

Search strategy

#22 #21 AND #15 AND #7

#21 #16 OR #17 OR #18 OR #19 OR #20

#20 TOPIC: Ultrasonic

#19 TOPIC: Echography

#18 TOPIC: Sonography

#17 TS = Ultrason*

#16 TOPIC: Ultrasound

#15 #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14

#14 TOPIC: Bechterew* disease

#13 TOPIC: Still* disease

#12 TOPIC: (Sicca NEAR/2 syndrome)

#11 TOPIC: (Sjogren* NEAR/2 syndrome)

#10 TOPIC: (Caplan* NEAR/2 syndrome)

#9 TOPIC: (Felty* NEAR/2 syndrome)

#8 TOPIC: (Rheumat* NEAR/5 (nodule or arthritis))

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

#7 #1 OR #2 OR #3 OR #4 OR #5 OR #6

#6 TOPIC: (Pharmacoeconomic* or price* or pricing)

#5 TOPIC: (Cost NEAR (estimate* OR variable OR unit))

#4 TOPIC: (Fiscal or funding or financial or finance)

#3 TOPIC: ((Low OR high) near cost*)

#2 TOPIC: (Cost* near/3 (analy* or benefit* or control or saving or illness* or sharing))

#1 TOPIC: Economic*

EconLit (via Ovid) (1961–2015)

Date searched: 22 October 2015.

As this database is dedicated to economic studies, no filters were used in the search.

Search strategy

1. (rheumat$adj5 (nodule or arthritis)).tw.
2. (felty$adj2 syndrome).tw.
3. (caplan$adj2 syndrome).tw.
4. (sjogren$adj2 syndrome).tw.
5. (sicca adj2 syndrome).tw.
6. still$disease.tw.
7. bechterew$disease.tw.
8. or/1–7
9. synovitis.tw.
10. ((Synovial or synovium) adj5 (inflam$or hypertrophy)).tw.
11. or/9–10
12. Ultrasound.tw.
13. Ultrason$.tw.
14. Sonography.tw.
15. Echography.tw.
16. Ultrasonic.tw.
17. or/12–16
18. 8 AND 11 AND 17

NHS Economic Evaluation Database (1968–2014) (via Wiley Online Library)

Searched as part of The Cochrane Library search [see *The Cochrane Library: Cochrane Database of Systematic Reviews (1996–2015), Cochrane Central Register of Controlled Trials (1898–2015), Health Technology Assessment database (1989–2015), Database of Abstracts of Reviews of Effects (1946–2014) and NHS Economic Evaluation Database (1968–2014) (via Wiley Online Library)* on page 101].

Parameter search

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid) (1946–2015)

Date searched: 20 October 2015.

Search strategy

1. Economics/
2. "costs and cost analysis"/
3. Cost-effectiveness analysis/
4. Cost–benefit analysis/
5. Cost control/
6. Cost savings/
7. Cost of illness/
8. Cost sharing/
9. "deductibles and coinsurance"/
10. Medical savings accounts/
11. Health care costs/
12. Direct service costs/
13. Drug costs/
14. Employer health costs/
15. Hospital costs/
16. Health expenditures/
17. Capital expenditures/
18. Value of life/
19. exp economics, hospital/
20. exp economics, medical/
21. Economics, nursing/
22. Economics, pharmaceutical/
23. exp "fees and charges"/
24. exp budgets/
25. (Low adj cost).mp.
26. (High adj cost).mp.
27. (Health?care adj cost$).mp.
28. (Fiscal or funding or financial or finance).tw.
29. (Cost adj estimate$).mp.
30. (Cost adj variable).mp.
31. (Unit adj cost$).mp.
32. (Economic$or pharmacoeconomic$or price$or pricing).tw.
33. or/1–32
34. exp Arthritis, Rheumatoid/
35. Rheumatoid arthritis.tw.
36. (Felty$adj2 syndrome).tw.
37. (Caplan$adj2 syndrome).tw.
38. (Rheumat$adj2 (nodule or arthritis)).tw.
39. (Sjogren$adj2 syndrome).tw.
40. (Sicca adj2 syndrome).tw.
41. Still$disease.tw.
42. Bechterew$disease.tw.
43. or/34–42
44. Biologic$.tw.
45. Anti-TNF.tw.
46. TNF-antagonist.tw.
47. TNF-inhibitor.tw.
48. (Abatacept or orencia).tw
49. (Adalimumab or humira or exemptia).tw.
50. (Certolizumab or CDP870 or cimzia).tw.
51. (Etanercept or enbrel).tw.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

52. (Golimumab or CNTO 148 or simponi).tw.
53. (Infliximab or remicade or remsima or inflectra).tw.
54. (Rituximab or rituxan or mabthera or zytux).tw.
55. (Tocilizumab or atlizumab or actemra or RoActemra).tw.
56. or/44–55
57. Taper$.tw.
58. Withdraw$.tw.
59. Discontinu$.tw.
60. Stop$.tw.
61. or/57–60
62. 33 and 43 and 56 and 61

Ongoing studies

The search of ClinicalTrials.gov identified some relevant studies that were ongoing or for which the results were not reported at the time of writing (March 2016) (*Table 19*).

TABLE 19 Ongoing studies identified from search of ClinicalTrials.gov (31 March 2016)

NCT number	Trial title
NCT01205854	Aiming for Remission in Rheumatoid Arthritis (RA) – the ARCTIC Trial (ARCTIC)
NCT02219347	Biomarkers of Remission in Rheumatoid Arthritis (BioRRA)
NCT01526434	Health-related Quality of Life and Patient-reported Outcomes in Rheumatoid Arthritis Patients Treated with Certolizumab Pegol (SONAR-12)
NCT02140229	Is Ultrasound Remission a Real Remission? Does Ultrasound Permit to Achieve and Maintain the Remission in Rheumatoid Arthritis Patients More Efficiently Than Clinical Scores? (REVECHO)
NCT02321930	Musculoskeletal Ultrasound Assessment of Therapeutic Response of Tofacitinib in Rheumatoid Arthritis Patients
NCT01717859	Musculoskeletal Ultrasound in Predicting Early Dose Titration with Tocilizumab (RASTS)
NCT01443364	Open Label Study to Assess the Predictability of Early Response to Certolizumab Pegol in Patients with Rheumatoid Arthritis (SPEED)
NCT02064400	Pilot Study of Ultrasound in Rheumatoid Arthritis
NCT00854243	Role of Greyscale and Power Doppler Sonography in Therapy Monitoring in Early Rheumatoid Arthritis (RA)
NCT02202837	Study with Etanercept Focusing on Remission and Predictability of Remission in Real Life Clinical Practice (REACH RA)
NCT02056184	Targeted Ultrasound in Rheumatoid Arthritis (TURA)
NCT01752309	The Predictive Value of Ultrasound in Early Rheumatoid Arthritis (EVA)
NCT01282528	Ultrasonographic Monitoring of Response to Infliximab in Patients with Rheumatoid Arthritis (ULTRA)
NCT00781989	Ultrasonography as a Biomarker in Early Rheumatoid Arthritis
NCT01602302	Ultrasound and Withdrawal of Biological DMARDs in Rheumatoid Arthritis (RA-BioStop)

Appendix 4 Excluded studies

Table 20 shows the studies excluded at full-text sift, with reasons for exclusion.

TABLE 20 Excluded studies

First author	Year of publication	Reason for exclusion
Alfredo Chávez-López[220]	2007	Non-English language
Andersen[221]	2014	No outcome data meeting review inclusion criteria
Andonopoulos[222]	1995	Study not about synovitis
Baan[223]	2011	No outcome data meeting review inclusion criteria
Bajaj[224]	2007	No clinical comparator
Boesen[225]	2008	No relevant outcomes
Brown[226]	2006	No outcome data meeting review inclusion criteria
Bruyn[227]	2009	No clinical comparator
Carotti[228]	2002	No outcome data meeting review inclusion criteria
Cheung[50]	2010	No outcome data meeting review inclusion criteria
D'agostino[62]	2016	No clinical comparator
Damjanov[229]	2012	No outcome data meeting review inclusion criteria
da Silva Chakr[230]	2015	No outcome data meeting review inclusion criteria
Dejaco[231]	2012	No outcome data meeting review inclusion criteria
DiFranco[232]	2015	No outcome data meeting review inclusion criteria
Dohn[233]	2011	No outcome data meeting review inclusion criteria
Epis[234]	2014	No outcome data meeting review inclusion criteria
Foltz[235]	2012	No clinical comparator
Freeston[236]	2008	No clinical comparator
Fukae[237]	2013	No clinical comparator
Fukae[238]	2014	No clinical comparator
Funck-Brentano[239]	2009	Not all patients have RA
Funck-Brentano[240]	2013	No clinical comparator
Geng[241]	2014	No outcome data meeting review inclusion criteria
Haavardsholm[110]	2009	No clinical comparator
Hameed[242]	2008	No outcome data meeting review inclusion criteria
Harman[243]	2015	No clinical comparator
Harman[244]	2015	No outcome data meeting review inclusion criteria
Hermann[245]	2003	No outcome data meeting review inclusion criteria
Hmamouchi[246]	2011	Study not about synovitis
Janta[247]	2013	No outcome data meeting review inclusion criteria
Kawashiri[248]	2014	No clinical comparator

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 20 Excluded studies (*continued*)

First author	Year of publication	Reason for exclusion
Kawashiri[249]	2015	No clinical comparator
Kelly[250]	2015	No outcome data meeting review inclusion criteria
Klauser[251]	2005	No clinical comparator
Klauser[252]	2010	Study not about synovitis
Krejza[253]	1998	Study about initial diagnosis of arthritis
Lillegraven[254]	2011	Study not about synovitis
Lillegraven[255]	2012	Not a US study
Makinen[256]	2007	Not a US study
Marks[26]	2015	No clinical comparator
Molenaar[48]	2004	Not a US study
Montoro[257]	2015	No outcome data meeting review inclusion criteria
Naredo[54]	2005	No outcome data meeting review inclusion criteria
Naredo[258]	2013	No outcome data meeting review inclusion criteria
Naredo[259]	2014	No outcome data meeting review inclusion criteria
Nordal[260]	2014	No outcome data meeting review inclusion criteria
Peluso[261]	2011	No clinical comparator
Ramirez[262]	2014	No outcome data meeting review inclusion criteria
Reiche[263]	2014	No clinical comparator
Schmidt[264]	2013	No relevant clinical comparator data
Spinella[265]	2012	No outcome data meeting review inclusion criteria
Stramare[266]	2014	No outcome data meeting review inclusion criteria
Strunk[267]	2013	No outcome data meeting review inclusion criteria
Szkudlarek[268]	2001	No separate data for RA patients
Taouli[269]	2004	Not a US study
Terslev[270]	2012	No outcome data meeting review inclusion criteria
Varsamidis[271]	2005	No clinical comparator
Wakefield[272]	2007	Study not about synovitis
Watanabe[273]	2012	No outcome data meeting review inclusion criteria
Yoshimi[274]	2015	No clinical comparator
Zheng[275]	2014	No outcome data meeting review inclusion criteria
Ziswiler[276]	2009	No outcome data meeting review inclusion criteria

Appendix 5 Data extraction tables

For calculations of diagnostic accuracy, US was counted as the reference standard and the accuracy of the clinical comparator was assessed using sensitivity (i.e. the proportion of TPs) and specificity (i.e. the proportion of TNs). Sensitivity is calculated as the number of TPs divided by the sum of the TPs and FNs; specificity is calculated as the number of TNs divided by the sum of the TNs and FPs.

TABLE 21 Data extraction table: Backhaus *et al.*[69]

First author (study name)	Backhaus[69]
Year	2013
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To determine the sensitivity to change of the US7 score among RA patients under various therapies and to analyse the effect of each therapeutic option over 1 year. To estimate predictors for the development of destructive bone changes* *p. 1163*
Population sample size	432
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA
Population baseline characteristics	81% female, 19% male Mean (SD, range) age 57 (12.8, 17–84) years Mean (SD, range) disease duration 8.3 (8.7, 0.08–58.3) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	73.2% prednisolone equivalent with a mean daily dosage of 8.8 mg per day Divided into four therapy groups at baseline: • group 1 – first-line DMARD after new initiation ($n = 118$; 27.3%) • group 2 – therapy switch from DMARD to a second DMARD ($n = 108$; 25.0%) • group 3 – first-line biologic after DMARD therapy ($n = 153$; 35.4%) • group 4 – therapy switch from biologic to a second biologic ($n = 53$; 12.3%).
Joints assessed	CE: DAS28 US: seven-joint count – wrist, MCP 2 and 3, PIP 2 and 3, MTP 2 and 5
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Settings for GSUS: frequency 16 MHz, length of scanner 40–42 mm Settings for PDUS: frequency 9.1 MHz, pulse repetition frequency 500–750 Hz (depending on machine setting). Machine NR Joints were evaluated for synovitis and tenosynovitis/paratenonitis and superficial bone erosions according to EULAR criteria and OMERACT definition including GSUS and PDUS GSUS was graded on the Scheel semiquantitative scale and PDUS was graded from 0 to 3 on the Szkudlarek semiquantitative scale
Who conducted US	NR
Comparator CE details	DAS28, ESR, CRP
Who conducted comparator CE	NR

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 21 Data extraction table: Backhaus *et al.*[69] (*continued*)

First author (study name)	Backhaus[69]
Follow-up duration (if relevant)	12 months
Primary outcome of study	Sensitivity to change of the US7 score among a large cohort of RA patients under various therapies (cDMARDs and/or biological therapy) and analysis of the effect of each therapeutic option over a period of 1 year
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	The US7 score was comparable to clinical and laboratory data, illustrating its potential to reflect the therapeutic response and sensitivity to change. Erosions declined significantly among patients who switched from one biologic to another, but were stable in the other groups

NR, not reported; SD, standard deviation.

TABLE 22 Data extraction table: Balsa *et al.*[105]

First author (study name)	Balsa[105]
Year	2010
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To investigate the accuracy of composite scores in classifying RA patients who are in remission using the absence of inflammatory activity detected by US as a gold standard* *p. 683*
Population sample size	97
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Patients were classified as being in clinical remission by their attending rheumatologist using subjective clinical judgement
Population baseline characteristics	70 female, 27 male Mean (SD, range) age 56 (12.2, 18–82) years Mean (SD, range) disease duration 5.9 (9.6, 1–18) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Taking DMARD therapy or biological agents at baseline. Excluded if taking a high dose of steroids (> 7.5 mg of prednisone daily) or had a history of intra-articular steroid joint injection during the past 6 months
Joints assessed	42 joints: PIP, MCP, wrist, elbow, bilateral glenohumeral, knee, ankle and midtarsal and MTP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Acuson Antares Siemens system with a linear probe (5–13 MHz) and a Doppler frequency of 5–8.9 MHz GSUS for synovial hypertrophy (SH) and/or joint effusion graded using a 0–3 semiquantitative scoring method (0 = no SH, 1 = mild SH, 2 = moderate SH and 3 = severe SH). PDUS graded using a 0–3 semiquantitative scoring method (0 = no PD signal, 1 = one or two vessels in small joints or up to three single vessels in large joints, 2 = less than half of the synovial area and 3 = more than half of the synovial area)
Who conducted US	Expert US rheumatologist
Comparator CE details	DAS28, SDAI
Who conducted comparator CE	Attending rheumatologist

TABLE 22 Data extraction table: Balsa *et al.*[105] (*continued*)

First author (study name)	Balsa[105]
Primary outcome of study	The relationship between clinical remission and imaging remission
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	'SDAI classification of remission is closer to the concept of an absence of inflammatory activity, as defined by the absence of a positive PD signal by US' (p. 683), than DAS28 classification of remission

SD, standard deviation.

TABLE 23 Data extraction table: Beckers *et al.*[106]

First author (study name)	Beckers[106]
Year	2004
Abstract or full paper	Full paper[a]
Study design	Diagnostic
Study objective	*To assess synovitis by ^{18}F-FDG PET in an individual joint analysis and in a global analysis of RA disease activity and to compare ^{18}F-FDG PET parameters with clinical, biological and sonographic (US) rheumatoid parameters* *p. 956*
Population sample size	21
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA, no DMARDs within 2 months
Population baseline characteristics	17 women, 4 men Mean (range) age 48 (34–69) years Mean (range) disease duration 11 (1–24) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Previously treated with 3.4 (range 1–8) DMARDs. Seventeen patients received low-dose oral corticosteroids (mean 7 mg/day, range 4–10 mg/day of prednisolone) and all received NSAIDs. None had taken a DMARD for 2 months before study entry
Joints assessed	356 joints in total: knees in all subjects and either wrists as well as MCP and PIP joints in 13 patients or ankles and the first MTP joints in the remaining eight patients
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS GSUS 13.0-MHz transducer and PDUS 5-MHz transducer (Aloka Prosound 5500). Pulse repetition frequency of 651 Hz for PDUS. Cut-off for US positivity was synovitis of ≥ 1-mm thick. PDUS scored on a 0–3 semiquantitative scale (0 = no signal, 1 = intermittent, 2 = persistent single spotting within the same location, 3 = multiple persistent spotting within the same location). Joints assessed at multiple sites (wrists and knees) were considered positive for GSUS and PDUS if at least one measurement/signal was identified
Who conducted US	One radiologist and one rheumatologist experienced in US
Comparator CE details	SJCs and TJCs
Who conducted comparator CE	Experienced study nurse

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 23 Data extraction table: Beckers *et al.*[106] (*continued*)

First author (study name)	Beckers[106]
Primary outcome of study	To compare ^{18}F-FDG PET parameters with clinical, biological and sonographic (US) rheumatoid parameters
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*^{18}F-FDG PET is a unique imaging technique that can assess the metabolic activity of synovitis and measure disease activity in RA* *p. 956*

^{18}F-FDG PET, 18F-fluorodeoxyglucose positron emission tomography.
a This research was originally published in JNM. Beckers *et al.* Assessment of disease activity in rheumatoid arthritis with F-18-FDG PET. *J Nucl Med* 2004;**45**(6):956–64. © by the Society of Nuclear Medicine and Molecular Imaging, Inc.[106]

TABLE 24 Data extraction table: Bhamra *et al.*[89]

First author (study name)	Bhamra[89]
Year	2014
Abstract or full paper	Abstract
Study design	Treatment
Study objective	*To evaluate the impact of clinic-based ultrasonography (MSUS) on the diagnosis and management of cases seen in [an] emergency rheumatology clinic* *p. 659*
Population sample size	17 RA (of 62 in study)
Population diagnosis of RA	NR
Population eligibility details (e.g. early RA, remission)	Patients referred to rheumatology clinic (not all RA)
Population baseline characteristics	NR for 17 RA patients. For 62 study patients, 25 men, 38 women; mean (range) age 57.17 (range 30–88) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	10 MCP and PIP joints, radiocarpal joint and ulnar styloid
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS GE Logiq E9 using a linear transducer. Scoring system NR
Who conducted US	Consultant rheumatologist
Comparator CE details	NR
Who conducted comparator CE	Referring clinician
Follow-up duration (if relevant)	NA
Primary outcome of study	Treatment decision change following US
Outcome(s) reported in main body of report	Treatment decision
Study authors' conclusions	*There is a positive impact of US in the rheumatology clinic; specifically highlighted multiple benefits in daily practice of reduced visits, discharge at first encounter and immediate management decisions* *p. 660*

NA, not applicable; NR, not reported.

TABLE 25 Data extraction table: Boyesen *et al.*[134]

First author (study name)	Boyesen[134]
Year	2011
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To examine the associations between modern imaging modalities and joint damage measured as 1-year MRI erosive progression in early RA patients* p. 176
Population sample size	84 recruited, 79 with 1-year follow-up data
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Early RA (< 1 year), treatment NR
Population baseline characteristics	65 patients (77%) female, 19 male Median (IQR) age 58 (47–67) years Median (IQR) disease duration 107 (70–188) days
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	CE: DAS28 US: dominant wrist
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS 8–16-MHz linear array transducer on a Diasus machine (Dynamic Imaging, Livingstone, UK). All findings were graded as 0 = none, 1 = mild, 2 = moderate or 3 = marked
Who conducted US	Trained user
Comparator CE details	DAS28-ESR
Who conducted comparator CE	NR
Follow-up duration (if relevant)	12 months
Primary outcome of study	Prediction of erosions at 1 year as measured by MRI
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*GSUS inflammation and MRI bone marrow oedema were independent predictors of MRI erosive progression in early RA patients on a group level. The exact prognosis of the individual patients could not be determined by imaging alone* p. 176

IQR, interquartile range; NR, not reported.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 26 Data extraction table: Brown *et al.*[135] and Ikeda *et al.*[136]

First author (study name)	Brown[135] and Ikeda[136]
Year	2008 and 2007
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To evaluate the long-term significance of subclinical synovitis and its relationship to structural outcome* p. 2958
Setting	UK, outpatient clinics
Population sample size	102 (90 with a full set of radiographs at both time points)[135] 107[136]
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	In remission (while taking cDMARDs – stable therapy for 6 months prior to baseline)
Population baseline characteristics	Mean age 57 (IQR 24–81) years 67% female; 33% male Duration of RA, median (range) 7 (2–38) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	cDMARDs
Joints assessed	Hand and wrist – MCP, radiocarpal, ulnar carpal, distal radioulnar, intercarpal compartments
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS ATL HDI 3000 machine (ATL Ultrasound, Bothell, WA, USA) with a 10–5 MHz linear array 'hockey-stick' transducer, according to the EULAR guidelines. Presence and location of synovial hypertrophy (SH) and erosions were recorded according to OMERACT definitions. SH was graded using a 0–3 semiquantitative scoring method (0 = no SH, 1 = mild SH, 2 = moderate SH and 3 = severe SH). PDUS images were scored using a 0–3 semiquantitative technique (0 = normal/minimal vascularity, 1 = mild hyperaemia, 2 = moderate hyperaemia and 3 = marked hyperaemia). Erosions were scored using a similar 0–3 semiquantitative scale, according to their location and severity/size
Who conducted US	A single experienced sonographer
Comparator CE details	Duration of morning stiffness; Likert scale and VAS for fatigue, joint pain, physician's assessment of disease activity and patient's global impression of health and disease activity; number of painful, tender and swollen joints as assessed by an independent trained metrologist; HAQ; RAQoL; ACR; DAS28; ESR; CRP
Who conducted comparator CE	Consultant rheumatologist and trained metrologist
Follow-up duration (if relevant)	12 months
Primary outcome of study	Progression of joint damage – erosions[135] Long-term radiological and clinical outcome[136]
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*The present study supports the use of sensitive imaging techniques for the accurate evaluation of disease status and the prediction of outcome in patients with RA, even when the findings of standard clinical measures of inflammatory activity have returned to normal* p. 2966

IQR, interquartile range.

TABLE 27 Data extraction table: Bugatti *et al.*[147] and Scirè *et al.*[137]

First author (study name)	[a]Bugatti[147] and Scirè[137]
Year	2012 and 2009
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To investigate whether baseline serum levels of the chemokine CXCL13 might predict clinical and ultrasonographic outcomes in patients with recent-onset RA* *Bugatti* et al.,[147] *p. 1* *To evaluate the usefulness of a systematic musculoskeletal ultrasonographic assessment in the detection of residual disease activity in patients with early RA who achieved clinical remission* *Scirè* et al.,[137] *p. 1092*
Population sample size	161 (155 at 12-month follow-up)[147] 106[137]
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Early RA; DMARD and glucocorticoid naive
Population baseline characteristics	From $n = 161$:[147] median (IQR) age 64 (50–73) years; 112 female (69.6%), 49 male; median (IQR) disease duration 3 (2–6) months From $n = 106$:[137] mean (SD) age 59.5 (14.4) years; 75 female, 31 male; mean (SD) disease duration 3.8 (2.8) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	$n = 72$ starting with MTX, $n = 31$ starting with HCQ, $n = 48$ starting with prednisone
Joints assessed	Bilateral shoulder, elbow, wrist (radiocarpal and midcarpal joint), MCP joints, PIP joints of the hands, sternoclavicular and acromioclavicular joints, knee, ankle and MTP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS GE Logiq 9 scanner (General Electrics Medical Systems, Milwaukee, WI, USA) with a multifrequency linear array transducer (10–15 MHz), performed according to EULAR guidelines GSUS and PDUS signals were scored on 0–3 semiquantitative scales.[277] GSUS scoring: 0 = normal, 1 = mild, 2 = moderate, 3 = marked. PDUS scoring: 0 = absence or minimal flow, 1 = mild (single-vessel signal), 2 = moderate (confluent vessels), 3 = marked (vessel signals in > 50% of the joint area)
Who conducted US	A single experienced operator
Comparator CE details	SJC and TJC on the 44-joint count, Ritchie Articular Index (RAI), global health assessment on a 0–100 mm VAS, evaluator global assessment of disease activity and patient global assessment of disease activity on a 0–10 cm VAS, ESR and CRP
Who conducted comparator CE	NR
Follow-up duration (if relevant)	12 months

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 27 Data extraction table: Bugatti *et al.*[147] and Scirè *et al.*[137] (*continued*)

First author (study name)	[a]Bugatti[147] and Scirè[137]
Primary outcome of study	Ultrasonic outcome (PDUS scores)[147] Residual disease activity[137]
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*CXCL13 is a promising prognostic marker in early RA, accurate in assessing the severity of synovitis and its persistence over time in response to conventional treatments* *Bugatti* et al.,[147] *p. 8* *The data support the specific role of US in detecting residual disease activity in early RA* *Scirè* et al.,[137] *p. 1096* *PD-positive synovial hypertrophy identifies ongoing inflammation even during remission and predicts short-term relapse* *Scirè* et al.,[137] *p. 1092*

IQR, interquartile range; NR, not reported; SD, standard deviation.
a Both studies included patients from the Early Arthritis Clinic (EAC) of the University Hospital of Pavia cohort.

TABLE 28 Data extraction table: Taylor *et al.*[100] and Cavet *et al.*[138]

First author (study name)	Taylor[100] and Cavet[138]
Year	2004 and 2009
Abstract or full paper	Full paper and abstract
Study design	Prognostic
Study objective	To investigate the ability of US and biomarkers to predict progressive joint damage
Population sample size	24
Population diagnosis of RA	ACR 1987 criteria[100]
Population eligibility details (e.g. early RA, remission)	RA patients who were followed in a 2-year blinded study comparing MTX + IFX with MTX alone in aggressive early RA RA, symptoms for 6 months to 3 years, a minimum of two swollen MCP joints despite treatment with oral MTX (minimum 8 weeks) and seropositivity for immunoglobulin M rheumatoid factor; either erosion of at least one MCP joint as demonstrated on plain radiography or GSUS or erosions of at least two MCP joints on GSUS and PDUS; stable dosage of 12.5–17.5 mg/week of folic acid at least 4 weeks prior to screening; if on corticosteroids must have been on a stable dose for 4 weeks (< 10 mg/day); screening laboratory tests[100]
Population baseline characteristics	18 female, 6 male[100] Mean (SD) age: MTX treated 51.4 (14.0) years, IFX + MTX treated 55.2 (11.8) years[100] Mean (SD) disease duration: MTX treated 1.64 (0.63) years, IFX + MTX treated 1.33 (0.64) years[100]

TABLE 28 Data extraction table: Taylor *et al.*[100] and Cavet *et al.*[138] (*continued*)

First author (study name)	Taylor[100] and Cavet[138]
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Treatment with MTX at a mean weekly dosage of approximately 15 mg for a mean duration of 0.91 years
Joints assessed	10 MCP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Scored for synovial thickening and for vascularity by PD area High-frequency (13-MHz) US and PD (14-MHz) imaging were performed using a 15L8 transducer (Acuson Sequoia, Siemens Medical Systems, Ultrasound Group, Issaquah, WA, USA) with constant settings in both GSUS and PDUS[100] GSUS images were evaluated for synovial thickness and assigned a score of 0–5. PDUS: number of colour Doppler pixels was determined in a defined region of interest for each joint and a total vascularity score was calculated as the sum of the individual joint scores[100]
Who conducted US	The same sonographer for all patients[100]
Comparator CE details	93 serum proteins associated with biological processes underlying joint damage were measured in serum samples TJC, SJC, morning stiffness duration (minutes), pain 0–10 VAS, patient and physician global assessment of disease activity 0–5 VAS, ACR 20/50/70 responses, DAS28, ESR
Who conducted comparator CE	An independent assessor
Follow-up duration (if relevant)	110 weeks
Primary outcome of study	Association of US and biomarkers with modified Sharp score at 110 weeks
Outcome(s) reported in main body of report	Prognostic and treatment
Study authors' conclusions	*Both ultrasonographic imaging and quantitative serum protein biomarkers can be used to estimate rates of progression and predict joint damage in RA. Serum proteins associated with change in TSS represent multiple biological pathways. Predictive models using US and biomarkers have the potential to improve patient outcomes* *Cavet* et al.[138] *p. 1464*

SD, standard deviation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 29 Data extraction table: Ceponis *et al.*[152]

First author (study name)	Ceponis[152]
Year	2014
Abstract or full paper	Full paper
Study design	Treatment decision and diagnostic
Study objective	*To investigate the usefulness of point-of-care hand and wrist joint US examination in patients with established RA* *p. 236*
Population sample size	51 (also included healthy control subjects, but data reported separately for RA patients)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA, no history or suspicion of fibromyalgia
Population baseline characteristics	Female 42 (91%), male 4 (9%) Mean (range) age 61.8 (28–82) years Mean (range) disease duration 16.6 (1–38) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Majority of patients were receiving cDMARDs, including MTX (60.9%; mean dosage 15.2 mg/week, range 5–25 mg/week) and leflunomide (15.2%; mean dosage 17.5 mg/day, range 10–20 mg/day) with or without a biologic agent [34.6%; either TNFi ($n = 14$), interleukin-6 inhibitor ($n = 1$) or RTX ($n = 1$)]
Joints assessed	MCP 1–5, PIP 2–5, wrist
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS A LOGIQ e US machine (GE Healthcare) equipped with a multifrequency 8- to 13-MHz linear transducer was used. Settings were standardised, with GSUS frequency of 12–13 MHz, gain of 58–64%, PD frequency of 6.7 MHz, gain of 9–12% and pulse repetition frequency of 0.6–0.8 Hz OMERACT definitions were used to assess joints for joint effusion and synovial hypertrophy, using a semiquantitative scale from 0 to 3 for GSUS and PDUS (0 = absence, 1 = mild, 2 = moderate and 3 = severe).[52,53] PD scoring: grade 0 = no intra-articular colour signal, grade 1 = single vessel signal(s), grade 2 = confluent colour signal in less than half of the intra-articular area and grade 3 = confluent colour signal in more than half of the intra-articular area[52]
Who conducted US	Experienced sonographer
Comparator CE details	CDAI, SJC, TJC, HAQ, pain VAS, morning stiffness, fatigue VAS, patient global assessment VAS
Who conducted comparator CE	Four board-certified rheumatologists
Follow-up duration (if relevant)	NA
Primary outcome of study	Agreement between US and clinical findings and its impact on physicians' confidence and clinical decision were assessed
Outcome(s) reported in main body of report	Treatment decision
Study authors' conclusions	*PD examination of the wrist and second/third MCP joints might be feasible and clinically meaningful in evaluation of disease activity in patients with established RA. US examination of the hand/wrist joints in RA increases physicians' confidence in their clinical decisions and can help to individualise DMARD and biologic agent use* *p. 236*

NA, not applicable.

TABLE 30 Data extraction table: Ciurtin *et al.*[90,158]

First author (study name)	Ciurtin[90,158]
Year	2013 and 2012
Abstract or full paper	Abstract
Study design	Treatment
Study objective	*To evaluate the usefulness of a 22 hand joints scoring system, adapted from the OMERACT recommendations, in assessing and differentiating patients with established RA from those with possible or definite early undifferentiated inflammatory arthritis. To establish the usefulness of the musculoskeletal US findings in guiding treatment decisions* *p. 295*
Population sample size	39 RA patients (of 98 in study)
Population diagnosis of RA	NR (established RA)
Population eligibility details (e.g. early RA, remission)	Patients referred to US service with inflammatory arthritis or suspected arthritis
Population baseline characteristics	NR for RA subset; for overall study: 69 females [mean (SD) age 50.03 (11.2) years) and 29 males [mean (SD) age 48.8 (11.7) years)
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	6/39 RA patients on biological therapy
Joints assessed	22 wrist joints and MCP and PIP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Machine NR. Scoring system adapted from the OMERACT recommendations; no details on scoring system reported
Who conducted US	Specialist US service
Comparator CE details	NR
Who conducted comparator CE	NR
Follow-up duration (if relevant)	NA
Primary outcome of study	Distinguishing RA from undifferentiated inflammatory arthritis and treatment decisions
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	In the 22 hand joints US scoring system, the presence of joint effusion grade 1 affecting < 5 joints and minimal synovial hypertrophy affecting < 3 joints did not correlate with any laboratory evidence of inflammatory or autoimmune abnormalities

NA, not applicable; NR, not reported; SD, standard deviation.

TABLE 31 Data extraction table: Dale *et al.*[153,154]

First author (study name)	Dale[153,154] (TaSER)
Year	2014 and 2013
Abstract or full paper	Full paper
Study design	Treatment
Study objective	*To determine the level of agreement and potential impact on DMARD escalation decisions and of adding musculoskeletal ultrasound (MSUS) assessment of disease activity to the DAS28* *Dale* et al.,[153] *p. 19*

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 31 Data extraction table: Dale *et al.*[153,154] (*continued*)

First author (study name)	Dale[153,154] (TaSER)
	To test whether the efficacy of DAS28-driven treat-to-target DMARD strategies could be improved by the addition of a regular MSUS *Dale* et al.,[154] *p. S338*
Population sample size	53 in the MSUS arm (110 with RA in the whole trial)
Population diagnosis of RA	ACR/EULAR 2010 criteria
Population eligibility details (e.g. early RA, remission)	Early RA/anti-citrullinated protein antibody-positive undifferentiated arthritis, untreated
Population baseline characteristics	Of the 53 patients in the MSUS arm: mean (SD) age NR; 30 females (59%); mean (SD) symptom duration 5.1 (2.8) months In the whole trial population ($n = 110$): median disease duration 4 months; DAS28 group contained a higher proportion of females (78%) than the MSUS group
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Step-up therapy (MTX; triple therapy; triple therapy with subcutaneous MTX; triple therapy + ETN)
Joints assessed	Dorsal recesses of 14 joints (second and third PIP joints, second and third MCP joints, wrist joints and second and fifth MTP joints bilaterally)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS All examinations were conducted using the same portable US machine (Voluson I, GE Healthcare) and a 10- to 16-MHz linear array probe (SP 10–16RS, GE Healthcare). PD examination was standardised using frequency high (machine preset), pulse repetition frequency 0.9 kHz, wall filter low and gain adjusted to below the level at which Doppler artefact appeared beneath bone. The presence of GS and PD synovitis positivity was graded on a 0–3 semiquantitative scale.[52] Active disease on MSUS was defined as the presence of a grade 1 or higher intra-articular PD signal in at least two joints. Therefore, the presence of a PD signal in two or more joints was used as a threshold for DMARD escalation
Who conducted US	Trained metrologist
Comparator CE details	DAS28, CRP, ESR, SJC, TJC, patient global assessment of disease activity, HAQ
Who conducted comparator CE	Trained metrologist
Follow-up duration (if relevant)	18 months
Primary outcome of study	Improvements in DAS44
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	*Compared with the DAS28, global RA disease activity assessment using a limited MSUS joint set provided additional disease activity information and led to altered treatment decisions in a significant minority of occasions* *Dale* et al.,[153] *p. 19* This may allow further tailoring of DMARD therapy by supporting DMARD escalation in patients with continuing subclinical synovitis and preventing escalation in symptomatic patients with minimal clinical and/or ultrasonographic synovitis Both groups exhibited similar, very robust improvements in clinical outcomes *MSUS disease activity assessment was not associated with improved clinical outcomes except for a higher rate of DAS44 remission after 18 months* *Dale* et al.,[154] *p. S339*

NR, not reported; SD, standard deviation.

TABLE 32 Data extraction table: Dougados *et al.*[139,149] and Cheung *et al.*[148]

First author (study name)	Dougados[139,149] and Cheung[148]
Year	2013 and 2014
Abstract or full paper	Full paper
Study design	Prognostic and management
Study objective	*To evaluate synovitis (clinical vs. ultrasound (US)) to predict structural progression in RA* *p. 665*
Population sample size	59 with data (from 77 recruited)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA, minimum disease activity of at least six swollen joints by CE, eligible for TNFi (investigator opinion)
Population baseline characteristics	Female 81%, male 19% Mean (SD) age 56 (12) years Mean (SD) disease duration 10 (8) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Mean number of previous cDMARDs 3.0. History of TNFis 32%. At baseline assigned to 4 months treatment with bDMARD (ETN, $n = 34$; ADA, $n = 23$, IFX, $n = 2$)
Joints assessed	MCP (×10), PIP (×10), wrist (×2) and MTP (×10) joints for US and CE
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Real-time scanners (e.g. Esaote Technos MPX, Esaote MyLab, Toshiba Aplio, Philips HD11, BK MiniFocus) using multifrequency linear transducers (7–12 MHz) Synovitis was defined according to the OMERACT definition, using a semiquantitative scale from 0 to 3 [GS: 0 = absence of synovial thickening, 1 = mild synovial thickening, 2 = moderate synovial thickening, 3 = marked synovial thickening; PD: 0 = absence of signal, no intra-articular flow, 1 = mild, one- or two-vessel signal (including one confluent vessel) for small joints and two to three signals for large joints (including two confluent vessels), 2 = moderate confluent vessels (> grade 1) and < 50% of normal area, 3 = marked vessel signals in more than half of the synovial area]
Who conducted US	Either a radiologist or a rheumatologist with experience
Comparator CE details	Clinical evaluation of synovitis and tender joints
Who conducted comparator CE	Either a rheumatologist or a research nurse with experience
Follow-up duration (if relevant)	2 years
Primary outcome of study	Association between structural deterioration and the presence of baseline synovitis, or its persistence, after 4 months of therapy
Outcome(s) reported in main body of report	Prognostic and treatment
Study authors' conclusions	*This study confirms the validity of synovitis for predicting subsequent structural deterioration, irrespective of the modality of examination of joints, but also suggests that both clinical and ultrasonographic examinations may be relevant to optimally evaluate the risk of subsequent structural deterioration* *p. 665*

SD, standard deviation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 33 Data extraction table: Ellegaard *et al.*[101]

First author (study name)	Ellegaard[101]
Year	2011
Abstract or full paper	Full paper
Study design	Treatment
Study objective	*To investigate the predictive ability of core outcomes applied in RA trials, including PDUS measurements differentiating patients who remain on TNFi therapy following 1 year* *p. 506*
Population sample size	109
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA, treated with TNFi (ADA, ETN or IFX)
Population baseline characteristics	78 females (71.1%) Mean (SD, range) age 57.9 (13.9, 25.5–84.3) years Mean (SD, range) duration of RA 10.4 (9.0, 1.0–34.6) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	TNFi (ADA, ETN or IFX)
Joints assessed	Wrists (most affected)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS Scanning was performed with a US machine (Siemens, Mountainview, CA, USA) using a linear array transducer with 14-MHz centre frequency, with no adjustments of Doppler parameters performed No details provided on scoring system used
Who conducted US	Head of the US unit, with 20 years' US experience; other investigators with several years of US training and 1 month of specific training on wrist scans
Comparator CE details	TJC, SJC, ESR, CRP, HAQ, patient general health VAS global, DAS28-CRP
Who conducted comparator CE	A rheumatologist unaware of the results of the US examination
Follow-up duration (if relevant)	1 year
Primary outcome of study	Remaining on TNFi therapy at 1 year
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	*There is now evidence to support that baseline PDUS, in contrast to clinical measures, can predict which patients will remain on TNFis 1 year after initiating therapy* *p. 506*

SD, standard deviation.

TABLE 34 Data extraction table: Filippucci *et al.*[107]

First author (study name)	Filippucci[107]
Year	2006
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To use PDUS to evaluate changes in synovial perfusion induced by adalimumab in the wrist joints of patients with RA* *p. 1433*
Setting	Two rheumatology centres in Italy
Population sample size	48 wrists of 24 patients
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active disease, aged ≥ 18 years, not pregnant, not in another trial of other biologic agents, no history of intra-articular steroid injections at the wrist, no changes in DMARD dose in the previous 3 months
Population baseline characteristics	18 female, 6 male Median (SD) age 61.5 (SD 10.5) years (IQR 48–67 years) Median (SD) disease duration 10 (SD 7.9) years (IQR 5–17 years)
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Started on ADA
Joints assessed	Wrists
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS AU5 Harmonic (Esaote Biomedica, Genoa, Italy) with a 10- to 14-MHz linear probe, standardised with a pulse repetition frequency of 1000 Hz and a colour mode frequency of 7 MHz Representative pictures of the highest expression of intra-articular PDUS signals were obtained. A score from 0 to 3 was assigned according to the overall expression of PDUS signals at the wrist level (semiquantitative visual scale: 0 = normal or minimal degree, 1 = mild degree, 2 = moderate degree and 3 = marked degree)
Who conducted US	An experienced operator blinded to both clinical and laboratory findings
Comparator CE details	Physician's global assessment of disease activity, ESR using the Westergren method, serum levels of CRP (upper reference level 4 mg/l; not measured at week 2)
Who conducted comparator CE	An experienced rheumatologist
Follow-up duration (if relevant)	12 weeks
Primary outcome of study	Change from baseline in clinical and US assessments
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	Our study suggests that: *PDUS is a feasible and sensitive imaging tool for assessing the response to treatment of synovitis at the small-joint level. In particular, we have shown that the wrist joint is a suitable anatomical site to be assessed by PDUS for detecting changes in synovial perfusion induced by systemic drug treatment* *p. 1437* Ongoing follow-up will add further insight into the persistence of considerable reductions in PDUS scores and their correlation with DAS28. In particular, long-term follow-up will provide information on the predictive value of rapid PDUS signal reduction for sustained remission of the disease at the small-joint level

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 34 Data extraction table: Filippucci *et al.*[107] (*continued*)

Joint swelling ($n = 24$, 192 joints)	PDUS negative	PDUS positive	Total
Clinically swollen	13	61	74
Clinical not swollen	32	86	118
Total	45	147	192
Joint tenderness ($n = 24$, 192 joints)	**PDUS negative**	**PDUS positive**	**Total**
Clinically tender	10	58	68
Clinically not tender	35	89	124
Total	45	147	192

Population	Diagnostic accuracy comparison	Sensitivity, %	Specificity, %	PPV, %	NPV, %
Joint swelling ($n = 24$, 192 joints)	CE with reference GSUS	41	71	82	27
Joint swelling ($n = 24$, 192 joints)	CE with reference GSUS	82	27	41	71
Joint tenderness ($n = 24$, 192 joints)	GSUS with reference CE	39	78	85	28
Joint tenderness ($n = 24$, 192 joints)	GSUS with reference CE	85	28	39	78

IQR, interquartile range; SD, standard deviation.

TABLE 35 Data extraction table: Gandjbakhch *et al.*[91]

First author (study name)	Gandjbakhch[91]
Year	2008
Abstract or full paper	Abstract
Study design	Treatment decision
Study objective	*To evaluate therapeutic decisions in daily practice and determine if US examination may influence therapeutic decisions in the management of RA patients* p. S467
Population sample size	52
Population diagnosis of RA	ACR criteria
Population eligibility details (e.g. early RA, remission)	RA patients referred for therapy adjustment
Population baseline characteristics	Sex NR Mean age 54 years Mean disease duration 10.3 years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	cDMARD, $n = 43$; TNFi, $n = 7$; no ongoing DMARDs, $n = 2$
Joints assessed	US: bilaterally on wrists, MCP 2–5 joints, PIP 2–5 joints, elbows, shoulders, knees and MTP 2–5 joints CE: DAS28
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Machine and scoring system NR

TABLE 35 Data extraction table: Gandjbakhch *et al.*[91] (*continued*)

First author (study name)	Gandjbakhch[91]
Who conducted US	An experienced rheumatologist
Comparator CE details	DAS28
Who conducted comparator CE	One of nine trained rheumatologists
Follow-up duration (if relevant)	NA
Primary outcome of study	Change in treatment decision (%) following US
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	*Rheumatologists' therapeutic decision was mainly made according to disease activity assessed by the DAS28. US findings increased clinician confidence and resulted in a change in therapeutic decision for 13% of the patients. US examination seems useful to detect residual inflammatory activity* p. S468

NA, not applicable; NR, not reported.

TABLE 36 Data extraction table: Garrigues *et al.*[108]

First author (study name)	Garrigues[108]
Year	2013
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To evaluate concordance between CE and US of joints in a heterogeneous group of patients with rheumatoid arthritis* p. 597
Population sample size	40
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Age ≥ 18 years
Population baseline characteristics	29 women and 11 men Mean (SD) age 55.9 (14) years Mean (SD) disease duration 11.2 (8.7) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Most patients receiving intravenous treatment for RA
Joints assessed	40 joints, namely the 28 joints included in the DAS28 (shoulders, elbows, wrists, MCP joints, PIP joints and knees) and tibiotalar and MTP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Multiplanar GSUS (B-mode) and PDUS images were obtained using commercially available real-time scanners (Esaote MyLab60 and Philips IU22) and multifrequency linear transducers (7–12.5 MHz) Synovitis defined according to 2005 OMERACT definitions GSUS: 0–3 semiquantitative scoring: 0 = no synovial thickening; 1 = mild synovial thickening (minimal synovial thickening not bulging beyond bone surfaces); 2 = moderate synovial thickening (synovial thickening bulging beyond bone surfaces without extension along the diaphysis); 3 = marked synovial thickening (synovial thickening bulging beyond the bone surfaces with extension along at least one of the diaphysis)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 36 Data extraction table: Garrigues *et al.*[108] (*continued*)

First author (study name)	Garrigues[108]
	PDUS: 0–3 semiquantitative scoring: 0 = no signal, no intra-articular flow; 1 = mild, until three isolated points or two confluents points or one confluent point and up to two isolated points; 2 = moderate vessel confluence (> grade 1) occupying < 50% of the normal synovial surface area; 3 = marked vessel confluence occupying > 50% of the normal synovial surface area
Who conducted US	Two sonographers (one radiologist, one rheumatologist)
Comparator CE details	SJC and TJC
Who conducted comparator CE	Rheumatologist
Primary outcome of study	Concordance between CE and US
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*Ultrasound adds information to clinical examination, most notably at the shoulders, wrists and MTP joints. Concordance was moderate to strong at other joint sites* p. 597

SD, standard deviation.

TABLE 37 Data extraction table: Gartner *et al.*[109]

First author (study name)	Gartner[109]
Year	2013
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To compare clinically active joints with sonographically active joints in patients with RA, applying different sonographic definitions of an active joint* p. 2005
Setting	Rheumatology outpatient clinic in Austria
Population sample size	90 patients (60 in clinical remission, 30 not in clinical remission)
Population diagnosis of RA	ACR 2010 criteria
Population eligibility details (e.g. early RA, remission)	Remission and active disease, no restriction on medication
Population baseline characteristics	In clinical remission: female 78.3%; mean (SD) age 60.1 (10.8) years; mean (SD) disease duration 9.4 (8.9) years; mean (SD) DAS28 2.2 (0.5) Not in clinical remission: female 76.7%; mean (SD) age 60.1 (11.3) years; mean (SD) disease duration 10.3 (7.7) years; mean DAS28 3.8 (1.1)
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	CE: 28 joints US: 11 joints of each hand, including PIP, MCP and wrist joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS and GSUS Logiq E9 (General Electric) with a ML 6–15 transducer, frequency range 9–15 MHz; for PD, pulse repetition frequency was set between 500 Hz and 800 Hz and receiver gain settings were controlled to eliminate the appearance of artefacts GS and PD signals for signs of synovitis were graded using a 0–3 semiquantitative scoring system (0 = none, 1 = mild, 2 = moderate and 3 = severe)[52]
Who conducted US	An experienced sonographer who had no access to the clinical and laboratory data and who was unaware of the results of the clinical joint examination

TABLE 37 Data extraction table: Gartner *et al.*[109] (*continued*)

First author (study name)	Gartner[109]
Comparator CE details	TJCs and SJCs were performed on 28 joints in all patients. Swelling and tenderness were defined in accordance with the standardised assessment recommendations from EULAR, whereby, in the presence of any doubt, a joint was considered 'not swollen'; 100-mm VAS for pain and global assessment of disease (patient's and evaluator's); duration of morning stiffness (minutes); HAQ-DI; CRP, ESR; CDAI; SDAI; DAS28
Who conducted comparator CE	NR
Primary outcome of study	Comparative analysis of sonographic and clinical joint assessment
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*. . . sonography as a tool to detect synovitis may have sufficient value only when the signals are high and that low signals may not necessarily represent inflammation and, in contrast to the clinical findings, are not related to disability* *p. 2012* Thus, the present findings reveal that more detailed assessment of sonographic data is needed to fully appreciate the value of US in the follow-up of patients with RA

NR, not reported; SD, standard deviation.

TABLE 38 Data extraction table: Haavardsholm and Ostergaard[110]

First author (study name)	Haavardsholm[110]
Year	2009
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To evaluate the responsiveness of MRI and US compared with conventional measures of disease activity and structural damage in patients with RA during the first year of treatment with TNFis* *p. 1572*
Population sample size	36
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA (started on TNFi treatment)
Population baseline characteristics	80.6% female Median (IQR) age 52.8 (24.0) years Median (IQR) disease duration 7.6 (8.0) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Started on TNFi
Joints assessed	One wrist for US, both hands and wrists for CE
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS All of the US measurements were performed using an 8- to 16-MHz linear array transducer (Diasus, Dynamic Imaging, Livingstone, UK) Synovitis/tenosynovitis and effusions were graded using a 0–4 semiquantitative scale (0 = none, 1 = uncertain, 2 = minimal, 3 = medium and 4 = high amount of hypoechoic material)
Who conducted US	An experienced ultrasonographer
Comparator CE details	28 SJC, 28 TJC, patient- and investigator-perceived pain 100-mm VAS, disability via modified HAQ, DAS28-ESR, SDAI, CDAI, CRP, ESR

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 38 Data extraction table: Haavardsholm and Ostergaard[110] (*continued*)

First author (study name)	Haavardsholm[110]
Who conducted comparator CE	A trained research nurse
Follow-up duration (if relevant)	12 months
Primary outcome of study	SRM
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*The most responsive measure of inflammation when evaluating TNFi medication was a composite measure comprising MRI synovitis, tenosynovitis and bone marrow oedema, and this may be a promising outcome measure in clinical studies* *p. 1572*

IQR, interquartile range.

TABLE 39 Data extraction table: Haarvardsholm *et al.*[79]

First author (study name)	Haarvardsholm[79] (ARCTIC)
Year	2015
Abstract or full paper	Abstract
Study design	Treatment
Study objective	To investigate tight control strategies to see whether or not a treatment strategy additionally including US would lead to better clinical and radiographic outcomes compared with a conventional treatment strategy
Population sample size	118 in the US arm (105 had completed the study at the time of abstract publication) 112 in the conventional strategy arm (99 had completed the study at the time of abstract publication)
Population diagnosis of RA	ACR/EULAR 2010 criteria
Population eligibility details (e.g. early RA, remission)	Early RA (symptom duration up to 2 years), DMARD naive, indication for DMARDs
Population baseline characteristics	NR
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	DMARD naive. Treatment strategy in trial: MTX, then triple cDMARDs, then bDMARD
Joints assessed	44 joints of DAS44; PDUS and GSUS assessed 32 joints on a 0–3 scale[111]
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS and GSUS assessed on a 0–3 scale.[111] Machine details NR
Who conducted US	Experienced sonographers
Comparator CE details	DAS, SJC
Who conducted comparator CE	NR
Follow-up duration (if relevant)	2 years
Primary outcome of study	Patients meeting all three assessments at 16, 20 and 24 months with a DAS of < 1.6, no swollen joints at 16, 20 and 24 months, and no progression in vdHSS (< 0.5 units) from 16 to 24 months
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	No statistically significant difference between the two strategies for the primary outcome

NR, not reported.

TABLE 40 Data extraction table: Hammer and Kvien[68] and Hammer *et al.*[111]

First author (study name)	Hammer[68,111]
Year	2011 and 2010
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To examine associations between ultrasonography assessments (GS and PD) of a large number of joints and traditional assessments of disease* *Hammer* et al.,[111] *p. 1349* *To explore the associations between a comprehensive ultrasonographic assessment of joints, tendons and bursae and previously described reduced joint counts (7-, 12-, 28- and 44- joint score)* *Hammer and Kvien*[68] *p. 1*
Population sample size	20
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA, commencing ADA
Population baseline characteristics	15 female (75%) Median (range) age 53 (21–78) years Median (range) disease duration 7.5 (1–26) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Starting ADA
Joints assessed	US: PIP 1–5 (dorsal), MCP 1–5 (dorsal), carpometacarpal 1–5 (dorsal), wrist (radiocarpal, intercarpal and radioulnar joints) (dorsal), elbow (anterior and posterior), shoulder (glenohumeral and acromioclavicular joints) (anterior, posterior and upper), hip (anterior), knee (anterior and lateral), ankle (talocrural joint) (anterior), four major foot joints (talonavicular, subtalar, calcaneocuboidal and cuneonavicular) (anterior and lateral), tarsometatarsal 1–5 (dorsal), MTP 1–5 (dorsal) and the interphalangeal (dorsal) joint of the first toe (a total of 78 joints) CE: PIP 1–5, MCP 1–5, wrist, elbow, shoulder, knee, ankle and MTP 1–5
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US examinations were performed with a 5- to 13-MHz probe and fixed settings optimal for PD signals (Siemens Antares, Sonoline; Siemens Medical Solutions, CA, USA). The same US machine and the same PD setting optimised for more superficial structures (most of the joints assessed) were used throughout the study All joints were scored according to OMERACT criteria for GSUS (presence of synovitis and joint fluid) and PDUS (presence of vascularisation) (0 = none; 1 = minor; 2 = moderate; 3 = major presence)
Who conducted US	One experienced sonographer
Comparator CE details	Tenderness and swelling of 40 joints (for comparisons between clinical and US examinations, a B-mode score of ≥ 1 was used to define a joint as inflamed), patient and study nurse global disease activity VAS, ESR, CRP, DAS28-ESR, SDAI, CDAI
Who conducted comparator CE	One of two study nurses, both with > 5 years' experience with joint counts in clinical studies
Follow-up duration (if relevant)	12 months
Primary outcome of study	Association between US and CE
Outcome(s) reported in main body of report	Diagnostic

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 40 Data extraction table: Hammer and Kvien[68] and Hammer *et al.*[111] (*continued*)

First author (study name)	Hammer[68,111]
Study authors' conclusions	*The comprehensive US assessments were associated with clinical and laboratory variables of disease activity and were highly sensitive to change during treatment with biological agents* *Hammer* et al.,[111] *p. 1349* *The reduced joint combinations were highly associated with the 78-joint score . . . [indicating] that an approach focusing on few joints and tendons gives equivalent information about the inflammatory activity in RA patients to a comprehensive US examination* *Hammer and Kvien*[68] *p. 1* The optimal combination of joints and tendons for a valid, reliable and feasible US measurement should be further explored to define a US score for follow-up of RA patients on biological treatment[68]

TABLE 41 Data extraction table: Hayashi *et al.*[92]

First author (study name)	Hayashi[92]
Year	2014
Abstract or full paper	Abstract
Study design	Diagnostic
Study objective	*To compare US findings with joint examination findings sorted by the presence of tenderness and/or swelling in the hand (proximal) interphalangeal (IP/PIP), MCP and wrist joints* *p. S181*
Population sample size	208
Population diagnosis of RA	NR
Population eligibility details (e.g. early RA, remission)	RA patients
Population baseline characteristics	158/208 female (76%) Mean age 66 years Mean disease duration NR
Population treatment (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	MCP, PIP, wrists
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US synovitis was defined as a GS imaging score of ≥ 1 (graded 0–3) or a synovial PD signal score of ≥ 2 (graded 0–3). Details of the machines and scoring systems used were not provided
Who conducted US	NR
Comparator CE details	Clinical joint assessments determined the presence of tenderness and/or swelling
Who conducted comparator CE	NR
Follow-up duration (if relevant)	NA
Primary outcome of study	Detection of synovitis by US or CE

TABLE 41 Data extraction table: Hayashi *et al.*[92] (*continued*)

First author (study name)	Hayashi[92]
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	Clinical joint examination of the IP/PIP joints overestimated, and of the wrist joints underestimated, synovitis, compared with US. The importance of US examination in daily clinical practice may differ among joint sites

NA, not applicable; NR, not reported.

TABLE 42 Data extraction table: Horikoshi *et al.*[112]

First author (study name)	Horikoshi[112]
Year	2010
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To compare MRI and US in the detection of joint inflammation in RA* *p. 556*
Population sample size	6
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	All patients were assessed clinically within 1 month of the study for disease activity using the DAS28-CRP
Population baseline characteristics	All females Mean (SD) age 50.2 (13.4) years Mean (SD) disease duration 13.5 (8.1) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Treated with one of the following bDMARDs: IFX, ETN, ADA or TCZ
Joints assessed	Intercarpal joints, radioulnar joints, second to fifth PIP joints and first to fifth MCP joints, first interphalangeal and radiocarpal joints. A total of 156 joints per patient were assessed by US
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US was performed using a Aplio SSA-700A system (Toshiba, Tokyo, Japan) with a 12-MHz linear array and a 12-MHz 'hockey-stick' array transducer. Joint inflammation on GSUS was defined as a hypoechoic intracapsular area and on PDUS was represented by the presence of positive findings of flow signal. These findings on both GSUS and PDUS were scored on a scale from 0 to 3[52]
Who conducted US	Two experienced rheumatologists
Comparator CE details	DAS28-CRP
Who conducted comparator CE	NR
Primary outcome of study	Agreement between US and MRI
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*Findings of PDUS correlated with those of MRI. Low-field MRI and PDUS are useful tools for the assessment of patients with RA* *p. 556*

NR, not reported; SD, standard deviation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 43 Data extraction table: Ikeda *et al.*[113,146]

First author (study name)	Ikeda[113,146]
Year	2013 and 2012
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To demonstrate that structural damage progression is associated with time-integrated PDUS signals more significantly than with time-integrated DAS28 in RA patients receiving MTX or biological agents* *Ikeda* et al.,[146] *p. 550*
Population sample size	57 ($n = 57$ with data at 24 weeks out of $n = 69$)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA, required treatment with MTX or bDMARD
Population baseline characteristics	40 female (73%), male 29 Mean (SD) age 54.9 (14.0) years Median (IQR) disease duration 37 (17–111) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	MTX 60/69 (87%) Corticosteroid 41/69 (59%) TNFi: ETN, $n = 10$, IFX, $n = 9$, ADA, $n = 9$
Joints assessed	28 joints of the DAS28 for US and CE
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Either LOGIQ 7 PRO (GE Healthcare), LOGIQ E9 (GE Healthcare), Vlamo (Toshiba Medica Systems Corporation) or HI VISION Avius (Hitachi Medical Corporation) GS and PD semiquantitative scores (0–3) (Naredo scoring). GS: 0 = absent; 1 = mild; 2 = moderate; 3 = marked. PD: 0 = absent (no synovial flow); 1 = mild (three or fewer isolated signals); 2 = moderate (more than three isolated signals or confluent signal in less than half of the synovial area); 3 = marked (signals in more than half of the synovial area)
Who conducted US	Two rheumatologists trained in US
Comparator CE details	28-joint SJC and TJC, CRP, VAS for patient and physician global assessments
Who conducted comparator CE	NR
Follow-up duration (if relevant)	24 weeks
Primary outcome of study	Correlation of PD score and joint damage progression
Outcome(s) reported in main body of report	Prognostic and diagnostic
Study authors' conclusions	*Synovial PD activity more accurately reflects active synovial inflammation that causes joint destruction than conventional measures in RA patients treated with MTX* *Ikeda* et al.,[113] *p. 1967* TNFis can inhibit short-term radiographic progression in the presence of active synovitis

TABLE 43 Data extraction table: Ikeda *et al.*[113,146] (*continued*)

Sensitivity to predict non-radiographic progression at 24 weeks (CIs NR)[113]					
Population	Baseline measure (optimum cut-off so different by treatment group)	Sensitivity, %	Specificity, %	PPV, %	NPV, %
57 ($n = 57$ with data at 24 weeks out of $n = 69$)	Total GS score cut-off point of < 62	56	57	69	43
MTX ($n = 16$)	Total GS score cut-off point of < 60	78	43	64	60
TNFi ($n = 24$)	Total GS score cut-off point of < 70	65	57	79	40
TCZ ($n = 17$)	Total GS score cut-off point of < 62	60	71	75	56
57 ($n = 57$ with data at 24 weeks out of $n = 69$)	Total PD score cut-off point of < 21	69	76	83	59
MTX ($n = 16$)	Total PD score cut-off point of < 20	89	86	86	89
TNFi ($n = 24$)	Total PD score cut-off point of < 21	65	71	85	46
TCZ ($n = 17$)	Total PD score cut-off point of < 18	70	86	88	67
57 ($n = 57$ with data at 24 weeks out of $n = 69$)	DAS28-CRP cut-off point of < 9.0	64	81	85	57
MTX ($n = 16$)	DAS28-CRP cut-off point of < 10.8	89	71	80	83
TNFi ($n = 24$)	DAS28-CRP cut-off point of < 11.9	88	71	88	71
TCZ ($n = 17$)	DAS28-CRP cut-off point of < 9.0	80	71	80	71

IQR, interquartile range; NR, not reported; SD, standard deviation.

TABLE 44 Data extraction table: Inanc *et al.*[93]

First author (study name)	Inanc[93]
Year	2014
Abstract or full paper	Abstract
Study design	Treatment prediction, prediction of response to bDMARDs by baseline US and clinical features
Study objective	*To investigate the ability of ultrasonographic parameters to predict which patients with RA will benefit from treatment with TNFi in terms of EULAR response* *p. 468*
Population sample size	43
Population diagnosis of RA	Either rheumatoid factor or anti-cyclic citrullinated peptide (CCP) antibody positive

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 44 Data extraction table: Inanc *et al.*[93] (*continued*)

First author (study name)	Inanc[93]
Population eligibility details (e.g. early RA, remission)	RA, bDMARD naive, starting TNFi
Population baseline characteristics	34 female, 9 male Mean (SD) age: responders ($n = 28$) 46 (11); non-responders ($n = 15$) 47 (11) Mean disease duration 8.0 ± 6.7 years Mean DAS28 of 5.4 ± 1.1
Population treatment (e.g. bDMARDS or cDMARDs)	Starting TNFi
Joints assessed	28 joints from DAS28
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US synovitis GS and PD signals were semiquantitatively graded from 0 to 3 (scoring system NR) using MyLab 70 US machine (Esaote, Italy)
Who conducted US	Experienced sonographer
Comparator CE details	TJS/SJC, DAS28, HAQ scores, ESR, CRP
Who conducted comparator CE	NR
Follow-up duration (if relevant)	12 months (with response data at 3 months)
Primary outcome of study	Response to TNFi related to baseline US
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	Despite similar clinical features, baseline PD scores, despite similar clinical features, can predict which patients will respond to TNFi therapy. Patients who respond at the third month are more likely to achieve low disease activity at 1 year. Ultrasonographic response to TNFi treatment can be achieved substantially in the first 3 months, beyond which changes in US scores are mostly non-significant
Outcome data not added to main report	Patients who responded at the third month were significantly more likely to achieve low disease activity ($p = 0.019$) or remission ($p = 0.008$) at 1 year. There was a significant decrease in mean PD and GS sum scores from baseline to the third month ($p < 0.001$ for both), but not between the third and sixth months (PD, $p = 0.68$; GS, $p = 0.77$)

NR, not reported; SD, standard deviation.

TABLE 45 Data extraction table: Iwamoto *et al.*[155]

First author (study name)	Iwamoto[155]
Year	2014
Abstract or full paper	Full paper
Study design	Prospective cohort study (treatment prediction, 6-month follow-up)
Study objective	*To determine whether the comprehensive US assessment of synovial inflammation predicts relapse after discontinuation of treatment with a biologic agent (TNFi or TCZ) in patients with RA in clinical remission*

p. 1576

TABLE 45 Data extraction table: Iwamoto *et al.*[155] (*continued*)

First author (study name)	Iwamoto[155]
Population sample size	42 ($n = 40$ with data at 6 months)
Population diagnosis of RA	ACR 1987 and ACR/EULAR 2010 criteria
Population eligibility details (e.g. early RA, remission)	RA, clinical remission (i.e. a DAS28 of < 2.6), on bDMARD and willing to discontinue bDMARD
Population baseline characteristics	33 female, 9 male Mean (SD) age 59.6 (12.8) years Mean (SD) disease duration 8.2 (6.7) years
Population treatment (e.g. bDMARDS or cDMARDs)	At baseline on bDMARD, discontinued at start of study. Baseline bDMARDS: IFX, $n = 17$; ETN, $n = 3$; ADA, $n = 6$; GOL, $n = 5$; CTZ, $n = 1$; TCZ, $n = 10$. Baseline median (IQR) MTX dose 8.0 (6.0–12.0) mg/week ($n = 38$). Baseline median (IQR) prednisolone dose 3.0 (2.0–5.0) mg/day ($n = 7$)
Joints assessed	134 synovial sites in 40 joints: finger, toe: interphalangeal, PIP (2–5), MCP (1–5), MTP (1–5) joints, flexor digitorum tendons; wrist: radiocarpal joint, intercarpal joints, distal radioulnar joint, compartment II/IV/VI of extensor tendons; elbow: humeroradial joint, humeroulnar joint, olecranon bursa; shoulder: glenohumeral joint, long head of biceps tendon, subacromion/subdeltoid bursae; knee: suprapatellar recess, femorotibial joint, popliteal bursa (Baker's cyst); ankle: tibiotalar joint, extensor tendons, flexor tendons, peroneal tendons
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	Systematic multiplanar GSUS and PDUS Aplio XG, Viamo (Toshiba Medical Systems), HI VISION Avius or HI VISION Ascendus (Hitachi Medical) systems Severity of US findings was graded semiquantitatively on a scale of 0–3. Each patient's total GS and PD scores were calculated by summing the corresponding scores of 40 joints
Who conducted US	Six rheumatologists trained for musculoskeletal US who were blinded to clinical information and laboratory data
Comparator CE details	DAS28, SJC, TJC, VAS physician and patient global assessment, HAQ-DI, ESR, CRP
Who conducted comparator CE	Nine rheumatologists who were blinded to the baseline US findings
Follow-up duration (if relevant)	6 months
Primary outcome of study	Relapse rates following bDMARD discontinuation and association with baseline US values. Relapse outcome defined as a DAS28 of > 3.2
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	*In RA patients in clinical remission receiving treatment with a biologic agent, residual synovial inflammation determined by comprehensive US assessment predicted relapse within a short time after discontinuation of the treatment* *p. 1576*

IQR, interquartile range; ROC, receiver operating characteristic; SD, standard deviation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 46 Data extraction table: Kamishima *et al.*[114]

First author (study name)	Kamishima[114]
Year	2011
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To compare quantitative MRI and PDUS with conventional measures of disease activity in RA patients treated with the anti-interleukin 6 receptor antibody TCZ in terms of responsiveness at a few months to disease activity and ability to predict structural damage at 1 year* *p. 745*
Population sample size	29
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA, treated with TCZ
Population baseline characteristics	2 males, 27 females Mean (range) age 61 (27–74) years Median (range) duration of symptoms 8 (1–34) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	TCZ
Joints assessed	Bilateral MCP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS A 13-MHz linear array transducer was used (Hitachi EUP-L34P, Hitachi, Tokyo, Japan). PD settings (75-dB dynamic range, medium persistence, medium frame rate, low wall filter, 1300-Hz pulse repetition frequency, medium vein flow optimisation, 1300-Hz speed velocity) were identical throughout the examinations Graded on a 0–3 semiquantitative scale (0 = absence of signal, 1 = single vessel dots, 2 = vessel dots over less than half of the synovial area, 3 = vessel dots over more than half of the synovial area)[277]
Who conducted US	One of the three ultrasonographers specialising in musculoskeletal US
Comparator CE details	DAS28-ESR, TJC, SJC, VAS, CRP, ESR
Who conducted comparator CE	One of the five rheumatologists
Follow-up duration (if relevant)	1 year (SRM at 5 months)
Primary outcome of study	Responsiveness to disease activity (SRM)
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*Conventional measures are responsive but less reflective of future bone destruction than image analysis . . . MRI of bone erosion and quantitative PDUS may be both responsive and predictive of structural damage progression at 1 year* *p. 753*

TABLE 47 Data extraction table: Kane *et al.*[102]

First author (study name)	Kane[102]
Year	2003
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	To compare US with CE: *in the detection of effusion, suprapatellar bursitis and Baker's cyst of the knee in RA to determine whether US provides additional clinical information* *p. 966*
Population sample size	22
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA
Population baseline characteristics	20 female, 2 male Mean (SD, range) age 50.2 (15.83, 25–79) years Mean (SD, range) disease duration 10.5 (7.8, 1.5–33) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	19 taking DMARDs (not specified if bDMARDs or cDMARDs)
Joints assessed	Knee
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS Real-time US was performed using an ATL (Seattle, WA, USA) HDI 3000 machine with L7- to 4-MHz and CL10- to 5-MHz probes Ultrasonographic assessment of joint effusion was recorded at each site. A knee effusion was present if hypoechoic fluid compressible by the transducer was found in either medial or lateral compartments of the knee Scoring system not reported
Who conducted US	An experienced rheumatologist with > 10 years' experience in musculoskeletal ultrasonography
Comparator CE details	Tenderness, swelling
Who conducted comparator CE	An experienced rheumatologist with > 6 years' clinical rheumatology practice
Primary outcome of study	Detection of effusion, suprapatellar bursitis and Baker's cyst of the knee
Outcome(s) reported in main body of report	Diagnosis
Study authors' conclusions	*US is more sensitive than CE in the detection of suprapatellar bursitis, knee effusion and Baker's cyst in RA. CE underestimates knee inflammation in RA* *p. 966*

SD, standard deviation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 48 Data extraction table: Kelly *et al.*[94]

First author (study name)	Kelly[94]
Year	2013
Abstract or full paper	Abstract
Study design	Management
Study objective	*To describe the impact of US use by rheumatologists on the diagnosis and management of RA patients in routine UK clinical practice compared with not using US* p. 101
Population sample size	109 with relevant data (of 258 in the study)
Population diagnosis of RA	NR
Population eligibility details (e.g. early RA, remission)	Patients aged > 18 years, newly referred to the rheumatology clinic with suspected inflammatory arthritis
Population baseline characteristics	NR for 109 patients with data relevant to this review All 258 patients: US group 31% male, non-US group 35% male; US group mean (SD) age 51.28 (15.75) years, non-US group mean (SD) age 53.12 (SD 17.34) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	NR
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	NR
Who conducted US	Consultant
Comparator CE details	NR
Who conducted comparator CE	Consultant
Follow-up duration (if relevant)	NA
Primary outcome of study	Impact of US use by rheumatologists on the diagnosis and management of RA patients in routine UK clinical practice compared with not using US
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	*Routine use of US in newly referred patients is associated with an earlier diagnosis and earlier DMARD initiation in patients with RA* p. 101

NA, not applicable; NR, not reported; SD, standard deviation.

TABLE 49 Data extraction table: Luengroongroj *et al.*[95]

First author (study name)	Luengroongroj[95]
Year	2015
Abstract or full paper	Abstract
Study design	Treatment
Study objective	*To evaluate the role of subclinical synovitis detected by musculoskeletal US in predicting disease relapse in remission RA* *p. 107*
Population sample size	32
Population diagnosis of RA	NR
Population eligibility details (e.g. early RA, remission)	Remission CDAI of < 2.8
Population baseline characteristics	NR
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	DMARDs, tapered during the study (not specified if bDMARDs or cDMARDs)
Joints assessed	NR
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS Total PD score No further details reported
Who conducted US	NR
Comparator CE details	DAS28-CRP, number of DMARDs
Who conducted comparator CE	NR
Follow-up duration (if relevant)	3 months
Primary outcome of study	Disease flare
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*PDUS assessment may indicate the likelihood of remaining in a disease remission stage. It seems to be safe to reduce the dose of [DMARD] dose . . . for a short period of time, especially when DMARDs tapering is urgent. However, closed monitoring for disease relapse is needed, especially in patients with subclinical synovitis* *p. 107*

NR, not reported.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 50 Data extraction table: Luukkainen *et al.*[115]

First author (study name)	Luukkainen[115]
Year	2003
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To assess the relationship between clinically detected swelling and effusion diagnosed by US in MTP and talocrural (TC) joints in patients with RA* *p. 632*
Population sample size	30
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA
Population baseline characteristics	24 females Mean (range) age 62 (36–80) years Mean (range) duration of RA 14 (0.2–27) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	MTP, talocrural (TC)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS US measurements were carried out using a Siemens Sonoline Prima apparatus with a 7.5-MHz transducer. Diagnosis of effusion was based on the method of Koski[278] The normal upper limit for MTP joints, 2.9 mm (mean + 2 SD)s, was evaluated according to the control group. If the anechogenic space was $\leq$ 2.9 mm it was regarded as normal and if it was $\geq$ 3.0 mm it was regarded as effusion. Correspondingly, the normal upper limit for TC joints, 3.0 mm (mean + 2 SDs), was taken from the control group. If it was $\leq$ 3.0 mm it was regarded as normal and if it was $\geq$ 3.1 mm it was regarded as effusion
Who conducted US	A doctor, no further details provided
Comparator CE details	Clinical assessment of MTP and TC joints by inspection and palpation according to the EULAR handbook.[279] Swelling was evaluated on a scale from 0 (normal) to 1 (swelling)
Who conducted comparator CE	A doctor (author, from department of rheumatology), no further details provided
Follow-up duration (if relevant)	NA
Primary outcome of study	Relationship between CE and US
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*These preliminary results showed poor agreement between the clinical assessment of swelling and effusion detected by US in MTP and TC joints. Thus, US may considerably improve the diagnosis of synovitis in patients with RA* *p. 632*

NA, not applicable; NR, not reported; SD, standard deviation.

TABLE 51 Data extraction table: Luukkainen and Sanila[103]

First author (study name)	Luukkainen[103]
Year	2005
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To compare the relationship between clinically detected swelling and effusion diagnosed by US in elbow joints in patients with RA* *p. 228*
Population sample size	50 (also included healthy control subjects, but data were reported separately for RA patients)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA
Population baseline characteristics	44 females Mean (range) age 56 (20–77) years Mean (range) duration of RA 11.9 (0.2–32) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	Glenohumeral (GH) joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS The diagnosis of effusion of GH joints by US was based on the method by Koski.[278] US measurements were carried out using a Siemens Sonoline Prima apparatus with a 7.5-MHz transducer US was scored using a scale of 0 (normal) or 1 (effusion)
Who conducted US	One doctor, no further details provided
Comparator CE details	Clinical assessment of GH joints was carried out using palpation according to the EULAR handbook.[279] The possible synovial swelling of the joints was evaluated using a scale of 0 (normal) or 1 (swelling)
Who conducted comparator CE	One doctor, no further details provided
Primary outcome of study	Relationship between CE and US
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	The results of this study indicate that: *. . . clinical assessment of swelling and evaluation of effusion by US in elbow joints in patients with RA show only fair agreement. Thus, US may improve the accuracy of the diagnosis of synovitis in many cases in these patients* *p. 228*

NR, not reported.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 52 Data extraction table: Luukkainen and Sanila[104]

First author (study name)	Luukkainen[104]
Year	2007
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To assess the relationship between swelling detected on physical examination and effusion diagnosed by ultrasonography in glenohumeral (GH) joints in patients with RA* *p. 865*
Population sample size	50 (also included healthy control subjects, but data were reported separately for RA patients)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA
Population baseline characteristics	40 females Mean (range) age 56.9 (20–77) years Mean (range) duration of RA 11.6 (0.2–30) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	GH joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS The diagnosis of effusion of GH joints by US was based on the method by Koski.[278] US measurements were carried out using a Siemens Sonoline Prima apparatus with a 7.5-MHz transducer US was scored using a scale of 0 (normal) or 1 (effusion)
Who conducted US	One doctor, no further details provided
Comparator CE details	Clinical assessment of GH joints was carried out using palpation according to the EULAR handbook.[279] The possible synovial swelling of the joints was evaluated on a scale of 0 (normal) or 1 (swelling)
Who conducted comparator CE	One doctor, no further details provided
Primary outcome of study	Relationship between CE and US
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*These results showed poor agreement between the clinical assessment of swelling and effusion detected by US in GH joints. Therefore, US may considerably improve the accuracy of diagnosis of effusion in GH joints* *p. 865*

TABLE 53 Data extraction table: Mamoto *et al.*[96]

First author (study name)	Mamoto[96]
Year	2013
Abstract or full paper	Abstract
Study design	Diagnostic
Study objective	*To determine the reliability of assessments of swollen joints by patients, physicians and US* *p. 710*
Population sample size	124
Population diagnosis of RA	NR
Population eligibility details (e.g. early RA, remission)	Active disease (with RA), no details reported
Population baseline characteristics	NR
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	Wrist, MCP, PIP bilaterally (22 joints per patient – 2728 joints in total)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Synovial hypertrophy was scored semiquantitatively using GS (score 0–3) and PD (score 0–3) signals. Swollen joints were defined as a GS score of ≥ 2. No further details reported
Who conducted US	A US examiner
Comparator CE details	DAS28, swollen joint assessment
Who conducted comparator CE	Attending physician and US examiner (and patient self-assessment)
Primary outcome of study	Sensitivity
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*US can assess swollen joints in patients with RA more sensitively than patients or attending physicians* *p. 710*

NR, not reported.

TABLE 54 Data extraction table: Mandl and Balint[116,117]

First author (study name)	Mandl[116,117]
Year	2013 and 2012
Abstract or full paper	Full paper
Study design	Ancillary study to RCT
Study objective	*To evaluate the intraobserver reliability, face validity and discriminant capacity of different global US scoring systems for measuring synovitis in RA* *Mandl and Balint,[117] p. 1272* *To evaluate the metrological properties of composite disease activity indices in RA, utilising information derived from clinical, GSUS and PDUS examinations and to assess the classification of patients according to disease activity using such indices* *Mandl and Balint,[116] p. 879*
Population sample size	62
Population diagnosis of RA	ACR 1987 criteria

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 54 Data extraction table: Mandl and Balint[116,117] (*continued*)

First author (study name)	Mandl[116,117]
Population eligibility details (e.g. early RA, remission)	Moderate RA (i.e. a DAS28 of > 3.2 and ≤ 5.1)
Population baseline characteristics	50 (80.6%) female Mean (SD) age 53.8 (13.2) years Mean (SD) disease duration 8.8 (7.7) months Median (IQR) disease duration 6.5 (11) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	cDMARDs (MTX) 30 patients receiving various cDMARDs and 32 receiving ETN + MTX
Joints assessed	DAS28 joints (bilateral shoulder, elbow, and wrist joints, bilateral MCP joints 1–5, bilateral PIP joints 1–5, and bilateral knee joints) plus bilateral MTP joints 1–5 and bilateral ankle and talonavicular joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Systematic multiplanar GSUS and PDUS examinations were carried out with commercially available real-time scanners (e.g. Esaote MyLab70 XVG, Esaote Technos MPX, General Electric Logiq 9) using multifrequency linear transducers (6–18 MHz). US scanning techniques, GS and PD machine settings and definitions of abnormality were standardised among investigators. Sonographers were allowed to modify machine settings (e.g. gain, pulse repetition frequency) on individual machines for the best-quality images, to appropriately score each image Synovitis was defined according to OMERACT definitions. Synovitis on GSUS was evaluated using a 0–3 semiquantitative scale[52] (0 = absence of synovial thickening, 1 = mild synovial thickening, 2 = moderate synovial thickening and 3 = marked synovial thickening). PD activity was evaluated using a 0–3 semiquantitative scale [0 = absence of signal, no intra-articular flow; 1 = mild, one or two vessels (including one confluent vessel) for small joints and two to three signals for large joints (including two confluent vessels); 2 = moderate confluent vessels (> grade 1) in < 50% of the synovium; 3 = marked vessel signals in > 50% of the synovium]
Who conducted US	Sonographers (no further details)
Comparator CE details	DAS28, TJC, SJC, SDAI, global disease activity patient and physician 100-mm VAS, HAQ-DI, ESR, CRP
Who conducted comparator CE	An investigator (no further details provided)
Primary outcome of study	Intraobserver reliability of composite synovitis scoring systems
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*GSUS and PDUS have better reliability than generally used clinical indices for evaluating synovitis in RA* *Mandl and Balint,[117] p. 1273* *Multimodal indices incorporating US and clinical data had similar metrological properties to their clinical counterparts; certain indices allowed for a significantly larger number of patients to be classified as having either high or moderate disease activity* *Mandl and Balint,[116] p. 879*

IQR, interquartile range; SD, standard deviation.

TABLE 55 Data extraction table: Naredo *et al.*[55]

First author (study name)	Naredo[55]
Year	2007
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To evaluate the sensitivity to change of PDUS assessment of joint inflammation and the predictive value of PDUS parameters in disease activity and radiological outcome in patients with early RA* p. 116
Population sample size	42
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Early RA, starting cDMARDs
Population baseline characteristics	31 female, 11 male Mean (SD, range) age 53.6 (14.1, 24–77) years Mean (SD, range) disease duration 6.8 (3.6, 1.5–12) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Starting cDMARDs at baseline: one cDMARD, $n = 41$; two cDMARDs, $n = 1$. Prior to study, 64.3% were taking oral corticosteroids and 85.7% were taking NSAIDs
Joints assessed	28 joints for CE and US: bilaterally glenohumeral, elbow, wrist, MCPs, PIPs, knees
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Real-time scanner (Logiq 500CL; General Electric Medical Systems, Kyunngi, Korea) using multifrequency linear array transducers (7–12 MHz). Joint synovitis was defined as the presence of intraarticular effusion and/or synovial hypertrophy. PD parameters were adjusted at the lowest permissible pulse repetition frequency (PRF) to maximise sensitivity, resulting in PRF ranging from 500 Hz to 1000 Hz. PD was graded on a 0–3 semiquantitative scale (0 = absence, no intra-articular flow; 1 = mild, single-vessel signal or isolated signals; 2 = moderate, confluent vessels; 3 = marked vessel signals in more than half of the intra-articular area)[54]
Who conducted US	Single rheumatologist, experienced, blinded
Comparator CE details	DAS28, SJC, TJC, pain VAS, patient global VAS, HAQ Radiographic assessment by vdHSS measuring erosions and joint space narrowing
Who conducted comparator CE	Single rheumatologist, experienced, blinded (independent of US conduction)
Follow-up duration (if relevant)	1 year
Primary outcome of study	Sensitivity to change of overall PDUS joint assessment and the predictive value of sequential PDUS parameters in clinical, functional and radiological outcomes
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*PDUS is a sensitive and reliable method for longitudinal assessment of inflammatory activity in early RA. PDUS findings may have a predictive value in disease activity and radiographic outcomes* p. 116

SD, standard deviation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 56 Data extraction table: Naredo *et al.*[140]

First author (study name)	Naredo[140]
Year	2008
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To evaluate the validity, responsiveness and predictive value of PDUS monitoring of response to tumour necrosis factor (TNF) blocking agents in RA* *p. 2248*
Population sample size	367 (278 with complete clinical, laboratory and PDUS data)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	All patients beginning therapy with a TNFi
Population baseline characteristics	Mean (SD) age 53.7 (12.3) years Mean (SD) disease duration 9.1 (8.2) years Of the 278 patients with complete clinical, laboratory and PDUS data: 227 female, 51 male; mean (SD) age 53.3 (12.2) years; mean (SD) disease duration 9.6 (8.2) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Starting TNFi therapy
Joints assessed	28 joints, including the left and right glenohumeral, elbow and wrist joints, MCP joints, PIP joints of the hands and knee joints, were assessed for tenderness and swelling
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS GSUS and PDUS examination was carried out with Logiq 5 Pro (GE Healthcare, Kyunnggi-do, Korea) scanners in 23 centres and Logiq 7 (GE Healthcare, Tokyo, Japan) scanners in two centres, using multifrequency linear array transducers (7–12 MHz). US scanning technique, GS and PD machine settings and definitions of abnormality were standardised. Synovitis, tenosynovitis and bursitis were defined according to the OMERACT definitions Synovial fluid and synovial hypertrophy were graded on a 0–3 semiquantitative scale (0 = absent; 1 = mild; 2 = moderate; 3 = marked). Synovial, tenosynovial and intrabursal blood flow at each joint was evaluated by PDUS. Pulse repetition frequencies were 500–750 Hz and colour gains were 18–30 dB The intra-articular, tenosynovial and intrabursal PD signal was graded on a 0–3 semiquantitative scale [0 = absent (no synovial flow); 1 = mild (three or fewer isolated signals); 2 = moderate (more than three isolated signals or confluent signal in less than half of the synovial area); 3 = marked (signals in more than half of the synovial area)] Modified US DAS was calculated for all patients at each visit, by replacing the SJC from the DAS28 with the count of joints with a synovial fluid, synovial hypertrophy or PD signal as determined by US (USDAS28 SF score, USDAS28 SH score and USDAS28 PD score, respectively)
Who conducted US	The same rheumatologist (one rheumatologist at 23 centres, two rheumatologists at two centres; all experienced in US), unaware of the clinical, laboratory and radiographic findings and not involved in treatment decisions
Comparator CE details	DAS28, TJC, SJC, patient-rated pain and disease activity on a 100-mm VAS, HAQ, CRP, ESR (rheumatoid factor)

TABLE 56 Data extraction table: Naredo *et al.*[140] (*continued*)

First author (study name)	Naredo[140]
Who conducted comparator CE	The same rheumatologist (one rheumatologist in 24 centres, two rheumatologists in one centre), who was blinded with regard to the PDUS and radiographic findings
Follow-up duration (if relevant)	12 months
Primary outcome of study	Validity and responsiveness of comprehensive PDUS assessment of synovitis to monitor response to anti-TNF therapy
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	Study results show: *. . . persistence of a synovial PD signal appears to . . . [predict] radiological progression in patients with established RA who are treated with anti-TNF agents* *p. 2554*

SD, standard deviation.

TABLE 57 Data extraction table: Naredo *et al.*[118]

First author (study name)	Naredo[118]
Year	2013
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To investigate the sensitivity for detecting subclinical synovitis of different reduced-joint US assessment models compared with a comprehensive US assessment in RA patients in clinical remission* *p. 512*
Population sample size	67
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA clinical remission assessed by their rheumatologist; MTX treatment for at least 2 years; neither disease flare nor changes in therapy, including corticosteroid and MTX doses, in the previous 6 months
Population baseline characteristics	50 female (74.6%), 17 male Mean (SD) age 60.3 (15.0) years Mean (SD) disease duration 7.5 (5.8) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	MTX treatment for at least 2 years
Joints assessed	Glenohumeral, elbow, wrist, second to fifth MCP, second to fifth PIP of the hands, hip (i.e. anterior recess), knee, ankle and second to fifth MTP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Real-time scanner (Mylab 70 XVG, Esaote) equipped with two multifrequency linear array transducers, a 6- to 18-MHz transducer for superficial areas and a 4- to 13-MHz transducer for deep areas GSUS synovial hypertrophy was scored on a 0–3 semiquantitative scale (0 = absent; 1 = mild; 2 = moderate; 3 = marked). The synovial PD signal was scored on a 0–3

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 57 Data extraction table: Naredo *et al.*[118] (*continued*)

First author (study name)	Naredo[118]
	semiquantitative scale [0 = absent (no synovial flow); 1 = mild (three or fewer PD signals); 2 = moderate (more than three PD signals in less than half of the synovial area); 3 = marked (more than three PD signals in more than half of the synovial area)]
Who conducted US	A rheumatologist experienced in musculoskeletal US
Comparator CE details	DAS28 and SDAI
Who conducted comparator CE	One investigator
Primary outcome of study	Detection of subclinical synovitis by different reduced-joint US assessment models compared with a comprehensive US assessment in RA patients in clinical remission
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*US assessment of the wrist, MCP, ankle and MTP joints can be highly sensitive for detecting residual B-mode and Doppler joint inflammation in RA patients* p. 512

67 patients	US negative, B-mode = 0	US positive, B-mode > 0	Total
CE positive, a DAS28 of > 2.6	0	26	26
CE negative, a DAS28 of < 2.6	5	36	41
Total	5	62	67
67 patients	**US negative, PD = 0**	**US positive, PD > 0**	**Total**
CE positive, a DAS28 of > 2.6	10	16	26
CE negative, a DAS28 of < 2.6	22	19	41
Total	32	35	67
67 patients	**US negative, B-mode = 0**	**US positive, B-mode > 0**	**Total**
CE positive, SDAI of > 3.3	1	44	45
CE negative, SDAI of < 3.3	4	18	22
Total	5	62	67
67 patients	**US negative, PD = 0**	**US positive, PD > 0**	**Total**
CE positive, SDAI of > 3.3	18	27	45
CE negative, SDAI of < 3.3	14	8	22
Total	32	35	67

Diagnostic accuracy comparison, CE with reference US	Sensitivity, %	Specificity, %	PPV, %	NPV, %
DAS28 of > 2.6 with B-mode US reference (> 0 on a scale of 0 to 3)	42	100	100	12
DAS28 of > 2.6 with PDUS reference (> 0 on a scale of 0 to 3)	46	69	62	54
SDAI of > 3.3 with B-mode US reference (> 0 on a scale of 0 to 3)	71	80	98	18
SDAI of > 3.3 with PDUS reference (> 0 on a scale of 0 to 3)	77	44	60	64
Diagnostic accuracy comparison, US with reference CE	**Sensitivity, %**	**Specificity, %**	**PPV, %**	**NPV, %**
B-mode US (> 0 on a scale of 0 to 3), reference DAS28 of > 2.6	100	12	42	100
PDUS (> 0 on a scale of 0 to 3), reference DAS28 of > 2.6	62	54	46	69
B-mode US (> 0 on a scale of 0 to 3), reference SDAI of > 3.3	98	18	71	80
PDUS (> 0 on a scale of 0 to 3), reference SDAI of > 3.3	60	64	77	44

SD, standard deviation.

TABLE 58 Data extraction table: Naredo *et al.*[156,157]

First author (study name)	Naredo[156,157]
Year	2014 and 2015
Abstract or full paper	Full paper
Study design	Treatment prediction
Study objective	*To investigate the predictive value of synovitis detected by Doppler US in relation to failed tapering of biologic therapy in RA patients in sustained clinical remission* *Naredo* et al.,[157] *p. 1408*
Population sample size	77 study completers (of 80 recruited)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA, stable bDMARDS in the previous 12 months, clinical remission by DAS28 or SDAI, no more than 5 mg/day of prednisone, no NSAIDs for more than 1 week and no corticosteroid injections in the previous 6 months
Population baseline characteristics	52 women, 25 men
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	All stable bDMARDS. Twenty-three (29.9%) patients were treated with ADA, 21 (27.3%) with ETN, 18 (23.4%) with IFX, 7 (9.1%) with TCZ, 6 (7.8%) with ABT and 2 (2.6%) with GOL
Joints assessed	42 joints (including hands and feet): glenohumeral, elbow, wrist, MCP 2–5, PIP 2–5, hip, knee, ankle and MTP 2–5 joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Real-time scanner (Mylab 70 XVG; Esaote, Genoa, Italy) equipped with two multifrequency linear array transducers, a 6- to 18-MHz transducer for superficial areas and a 4- to 13-MHz transducer for deep areas, assessed by presence and grade according to OMERACT. A global index for GSUS synovitis and a global index for Doppler synovitis were calculated Synovial hypertrophy was scored on a 0–3 semiquantitative scale (0 = absent; 1 = mild; 2 = moderate; 3 = marked). The synovial PD signal was also scored on a 0–3 semiquantitative scale (0 = no synovial PD signal; 1 = mild (three or fewer PD signals within the synovial hypertrophy); 2 = moderate (more than three PD signals in less than half of the synovial hypertrophy); 3 = marked (PD signals in more than half of the synovial hypertrophy)
Who conducted US	Rheumatologist experienced in US
Comparator CE details	DAS28, SDAI, clinical and laboratory assessment every 3 months
Who conducted comparator CE	Two study investigators
Follow-up duration (if relevant)	12 months
Primary outcome of study	The predictive value of synovitis detected by Doppler US in relation to failed tapering of biologic therapy at 6 and 12 months in RA patients in sustained clinical remission
Outcome(s) reported in main body of report	Treatment
Study authors' conclusions	*The presence of Doppler-detected synovitis may predict biologic therapy tapering failure in RA patients in sustained clinical remission* *Naredo* et al.,[157] *p. 1408*

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 59 Data extraction table: Osipyants *et al.*[97,150]

First author (study name)	Osipyants[97,150]
Year	2013
Abstract or full paper	Abstract
Study design	Prognostic
Study objective	*To assess the significance of residual inflammation according to the presence of SJC and PD signals in relation to radiographic progression in patients receiving TCZ* *Osipyants* et al.,[150] *p. A755*
Population sample size	36
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA patients receiving TCZ
Population baseline characteristics	26 (72%) female, 10 male Median (range) age 51 (43–57) years Median (range) disease duration 54 (24–96) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	All patients receiving TCZ
Joints assessed	Bilateral, the most affected joints: wrist, MCP 2 and 3 and PIP 2 and 3 joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS 'Voluson-i' (GE, USA) with transducer (4–13 MHz) Each joint was scored according to the OMERACT definitions of pathology. Dichotomised into 'non- or low-active' (PD $\leq$ 1) and 'middle- or high-active' (PD > 1)
Who conducted US	Single operator
Comparator CE details	SJC, dichotomised into 'non- or low-active' (SJC $\leq$ 1) and 'middle- or high-active' (SJC > 1)
Who conducted comparator CE	NR
Follow-up duration (if relevant)	1 year
Primary outcome of study	Relation of US and clinical baseline parameters with radiographic progression at 1 year. One-year radiographic progression of the hands and feet was defined when the change in the vdHSS (change in TSS) was > 0.5 units per year
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*In TCZ-treated patients the US imaging activity score appears to be more predictive of radiographic progression than SJC status* *Osipyants* et al.,[150] *p. A755*

NR, not reported.

TABLE 60 Data extraction table: Pereira *et al.*[119]

First author (study name)	Pereira[119]
Year	2015
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To determine if there is a correlation between intra-articular PDUS and pain symptoms in RA* *p. 1975*
Population sample size	72 considered in two groups: painful MCP, $n = 34$; painless MCP, $n = 38$
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Established RA, chronic swelling of at least one MCP joint for at least 3 consecutive months and stable use of DMARDs for the previous 3 months. For painful group, pain of ≥ 4 on VAS 0–10. No other concomitant inflammatory diseases, no history of hand surgery or presence of severe hand deformities and no uncontrolled comorbidities
Population baseline characteristics	Female, *n* (%): painless group 35 (92); painful group 33 (97) Mean (range) age: painless group 60.3 (32–81) years; painful group 56.9 (29–88) years Mean (range) RA duration: painless group 17.3 (2–46) years; painful group 14.7 (2–38) years
Population treatment (e.g. bDMARDS or cDMARDs)	Corticosteroid use, *n* (%): painless group 15 (39.5), painful group 17 (50); DMARD monotherapy, *n* (%): painless group 17 (44.7), painful group 6 (17.6); DMARDs association, *n* (%): painless group 9 (23.7), painful group 11 (32.4); dose of MTX, mean (range): painless group 9.5 (0–25) mg, painful group 12.5 (0–25) mg: biologic therapy use, *n* (%): painless group 10 (26.3), painful group 15 (44.1); DMARD + biologic therapy use, *n* (%): painless group 7 (18.4), painful group 13 (38.2)
Joints assessed	MCP 2–5, bilaterally
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS MyLAb 60 (Esaote SpA, Genoa, Italy) and MyLab Twice (Esaote SpA, Genoa, Italy) equipped with 6- to 18-MHz broadband multifrequency linear transducer. MCP joints scanned using a multiplanar technique adopting the indications provided by EULAR guidelines. PD settings were standardised at pulse repetition frequency of 500–750 Hz and Doppler frequency between 9.1 and 11.1 MHz. OMERACT preliminary definitions were adopted Joints were examined by US according to GS, PD, erosion (present or absent) and 0–3 semiquantitative scores. PD: semiquantitative score: 0 = no intra-articular flow, 1 = single-vessel signal, 2 = confluent vessels and 3 = vessel signal in > 50% of the intra-articular area.[54] GS: semiquantitaive score: 0 = no synovial thickening, 1 = minimal synovial thickening, 2 = moderate synovial thickening with capsule distension and 3 = synovial thickening, extending to bone diaphysis[52]
Who conducted US	Two experienced rheumatology sonographers trained for the assessment of RA
Comparator CE details	CE (on same day as US), SJC
Who conducted comparator CE	Expert rheumatologist

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 60 Data extraction table: Pereira *et al.*[119] (*continued*)

First author (study name)	Pereira[119]
Follow-up duration (if relevant)	NA
Primary outcome of study	Correlation between intra-articular PD and pain symptoms
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*Intra-articular PD was not correlated with pain symptom in this study* *p. 1975*

NA, not applicable.

TABLE 61 Data extraction table: Ramirez García *et al.*[98]

First author (study name)	Ramirez García[98]
Year	2014
Abstract or full paper	Abstract
Study design	Prognostic, prospective cohort study
Study objective	*To analyse clinical and sonographic predictors of clinical flares in patients with RA in clinical remission* *p. 888*
Population sample size	28
Population diagnosis of RA	Criteria NR ('diagnosed with rheumatoid arthritis')
Population eligibility details	RA, clinical remission (defined as a DAS28 of < 2.6)
Population baseline characteristics	23 female, 5 male Mean (SD) age 51.9 (11.9) years Mean (SD) disease duration 108.2 (SD 92.6) months
Population treatment	75% on at least one cDMARD; 39.3% on biological treatment
Joints assessed	Knees and hands (wrists, MCP, PIP, flexor and extensor tendons)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS High sensitivity equipment (Acuson AntaresTM; Siemens AG, Erlangen, Germany) with a 8-to 12-MHz linear probe. Synovial hypertrophy (grades 0–3) and PD (grades 0–3) quantified
Who conducted US	Rheumatologist experienced in musculoskeletal US
Comparator CE details	Serum concentrations of several cytokines and angiogenic factors were analysed by enzyme-linked immunosorbent assay (ELISA) at baseline [results reported for ESR, concentration of transforming growth factor-beta1 (TGF-β1)]
Who conducted comparator CE	Analysis by RayBiotech Inc.
Follow-up duration	12 months
Primary outcome of study	Clinical and sonographic predictors of clinical flares
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*Although an elevated ESR or a low concentration of TGF-β1 at baseline seems to be associated with clinical reactivation of patients with RA in remission, the PD signal is the best predictor of disease reactivation at 12 months* *p. 888*

SD, standard deviation.

TABLE 62 Data extraction table: Reynolds *et al.*[141] and Rees *et al.*[151]

First author (study name)	Reynolds[141] and Rees[151]
Year	2009 and 2007
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To determine whether a range of single time point US measures of synovial disease and serological characteristics are able to predict progression of US-defined erosive disease in patients with established RA* *Reynolds* et al.,[141] *p. 473*
Population sample size	40 (25 with follow-up data)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Established RA
Population baseline characteristics	29 females Median (range) age 59 (34–84) years Median (range) disease duration 6 (1–29) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Of the 25 patients followed up, 3 were on anti-inflammatory drugs only, 14 were on DMARDs and 8 were treated with a TNFi (or started during the study)
Joints assessed	One PIP or MCP joint from each patient was chosen as representative of one of four categories of joint based solely on the clinical signs. The four categories were (1) not swollen or tender, (2) swollen only, (3) tender only and (4) swollen and tender
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Philips HDI 5000 (Philips Medical Systems, Andover, MA, USA) with a C7- to 15-MHz 'hockey-stick' transducer Erosions were graded on a 0–3 semiquantitative scoring system (Lund) (0 = absence of erosions; 1 = one to two erosions; 2 = more than two erosions; 3 = areas of regional bone destruction). GS images were graded on a 0–3 semiquantitative scoring system (0 = absence of synovial hypertrophy; 1 = small degree of synovial hypertrophy; 2 = moderate synovial hypertrophy; 3 = marked synovial hypertrophy). PD images were graded on a 0–3 semiquantitative scoring system (0 = sbsence of PD signal; 1 = single-vessel dots; 2 = confluent vessel dots over less than half of the area of the synovium; 3 = confluent vessel dots over more than half of the area of the synovium)
Who conducted US	Two radiologists
Comparator CE details	Swelling, tenderness, ESR (rheumatoid factor, anti-citrullinated protein antibody)
Who conducted comparator CE	Baseline: experienced rheumatologist; follow-up: NR
Follow-up duration (if relevant)	2 years (mean 26.8 months, range 24–32 months)
Primary outcome of study	Joint erosions
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*The majority of single time point US measures of synovial disease were not able to identify MCP or interphalangeal joints destined to develop progressive US-determined bone damage in patients with established RA* *Reynolds* et al.,[141] *p. 473*

NR, not reported.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 63 Data extraction table: Ribbens *et al.*[120]

First author (study name)	Ribbens[120]
Year	2003
Abstract or full paper	Full paper
Study design	Diagnostic and response to treatment
Study objective	*To evaluate using GSUS and PDUS and clinical assessment the response of hand joint synovitis in patients with active RA to treatment with TNFi (IFX)* *p. 562*
Population sample size	11
Population diagnosis of RA	ACR 1987 revised criteria
Population eligibility details (e.g. early RA, remission)	Active RA despite MTX, six or more swollen joints, 10 or more tender joints and one of the following: morning stiffness for > 45 minutes, ESR of > 28 mm/hour or CRP of > 20 mg/l. Also, presence of erosions
Population baseline characteristics	7 women, 4 men Mean (range) age 54 (25–74) years Mean (range) disease duration 9 (2–31) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Previously treated with a mean of 3.2 (range 1–5) cDMARDs. After baseline examination, treated with IFX and MTX
Joints assessed	Wrist, MCP and PIP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US (Aloka Prosound 5500; Aloka, Tokyo, Japan) was performed using a B-mode 13.0-MHz transducer and a PD 5-MHz pulse repetition frequency of 521 Hz. PD settings were standardised to a pulse repetition frequency of 651 Hz. Synovitis was defined as a hypoechoic or anechoic area in the joint space. Cut-off for US positivity was defined as a synovial thickness of ≥ 1 mm Doppler US measurements were simultaneously carried out and positive signals were scored according to a 0–3 semiquantitative scale (0 = no perfusion; 1 = low perfusion; 2 = moderate perfusion; 3 = intense intra-articular joint perfusion)
Who conducted US	One radiologist and one rheumatologist, experienced in US, both examined each patient
Comparator CE details	DAS28-ESR
Who conducted comparator CE	An independent physician experienced in joint assessment
Follow-up duration (if relevant)	6 weeks
Primary outcome of study	The response of hand (i.e. wrist, MCP and PIP) joint synovitis in patients with active RA to treatment with IFX
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*US is a feasible imaging modality for the measurement of the response of RA small-joint synovitis to therapy* *p. 562*

Additional data[120]

Wrist (20 joints)	GSUS negative	GSUS positive	Total
Clinically swollen	2	13	15
Clinically not swollen	3	2	5
Total	5	15	20
MCP (110 joints)	**GSUS negative**	**GSUS positive**	**Total**
Clinically swollen	17	47	64
Clinically not swollen	19	27	46
Total	36	74	110

TABLE 63 Data extraction table: Ribbens *et al.*[120] (*continued*)

PIP (103 joints)		GSUS negative	GSUS positive	Total	
Clinically swollen		14	25	39	
Clinically not swollen		44	20	64	
Total		58	45	103	
Population	**Diagnostic accuracy comparison**	**Sensitivity, %**	**Specificity, %**	**PPV, %**	**NPV, %**
Wrist	CE with reference US	87	60	87	60
MCP	CE with reference US	64	53	73	41
PIP	CE with reference US	56	76	64	69
Population	**Diagnostic accuracy comparison**	**Sensitivity, %**	**Specificity, %**	**PPV, %**	**NPV, %**
Wrist	US with reference CE	87	60	87	60
MCP	US with reference CE	73	41	64	53
PIP	US with reference CE	64	69	56	76

TABLE 64 Data extraction table: Riente *et al.*[121]

First author (study name)	Riente[121]
Year	2010
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To investigate the prevalence of US pathogenic abnormalities and to compare them with the clinical findings in the knee of RA patients* *p. 300*
Population sample size	100
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Patients with RA, no previous knee joint surgery and no corticosteroid injections within the previous 3 months
Population baseline characteristics	79 female, 21 male Mean (SD, range) age 58 (5.74, 22–82) years Mean (SD, range) disease duration 96 (70, 12–288) months
Population treatment (e.g. bDMARDS or cDMARDs)	200 knee joints
Joints assessed	Knee
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Logiq 9 (General Electrics Medical Systems, Milwaukee, WI, USA) with a linear probe operating at 10 MHz when studying joints and 14 MHz when studying tendons used in all centres. When synovial proliferation was detected, PD examination (pulse repetition frequency 500 Hz, Doppler frequency 7.5 MHz and Doppler gain to avoid random noise visualisation) was performed

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 64 Data extraction table: Riente *et al.*[121] (*continued*)

First author (study name)	Riente[121]
	Joint effusion, synovial hypertrophy and enthesopathy and bone erosions diagnosed according to OMERACT definitions. No details of scoring system reported
Who conducted US	One rheumatologist experienced in musculoskeletal US in each unit
Comparator CE details	Pain, tenderness and swelling (presence/absence)
Who conducted comparator CE	A rheumatologist (not involved in US examination)
Primary outcome of study	Prevalence of US pathogenic abnormalities
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*US examination of the knee is more sensitive than CE in the detection of joint inflammation and allows for the identification of different patterns of pathological changes at the knee level* *p. 300*

SD, standard deviation.

TABLE 65 Data extraction table: Riente *et al.*[122]

First author (study name)	Riente[122]
Year	2011
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To investigate the prevalence of US abnormalities in the foot of RA patients and compare these with clinical findings* *p. 1*
Population sample size	100
Population diagnosis of RA	ACR 1987 revised classification criteria
Population eligibility details (e.g. early RA, remission)	RA diagnosed, attending outpatient or inpatient clinic, no prior joint surgery, no corticosteroid injection to foot within 3 months
Population baseline characteristics	72 female, 28 male Mean (SD) age 56 years (4.8) Mean (SD) disease duration 65 (75) months
Population treatment (e.g. bDMARDS or cDMARDs)	NR (apart from no steroid injection within 3 months prior to study)
Joints assessed	Bilateral MTP (2–5), PIP (2–5) and midfoot joints [taloavicular, calcaneocuboid, medial, intermediate and lateral navicular-cuneiform and cuneiform-metatarsal joints and cuboid (fourth) and metatarsal (fifth) joints]
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US examinations used a Logiq9 (General Electrics Medical Systems, Milwaukee, WI, USA) system with linear probe at 14 MHz, and a MyLab70XVG (Esaote SpA, Genoa, Italy) system with multifrequency linear probe at 16 MHz. Multiplanar scanning technique was according to EULAR guidelines (dorsal and plantar aspects). Diagnosis was according to preliminarly definitions from OMERACT, using a semiquantitative grading method (0–3)
Who conducted US	Rheumatologist experienced in US
Comparator CE details	CE for swelling, pain and tenderness

TABLE 65 Data extraction table: Riente *et al.*[122] (*continued*)

First author (study name)	Riente[122]
Who conducted comparator CE	Rheumatologist not involved in the US examination
Primary outcome of study	Prevalence of US abnormalities in the foot
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*US examination appears to be a useful technique to study foot joint and tendon involvement in RA patients . . . [and] is more sensitive than CE to detect joint inflammation* *p. 1*

NR, not reported; SD, standard deviation.

TABLE 66 Data extraction table: Salaffi *et al.*[160]

First author (study name)	Salaffi[160]
Year	2008
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To assess the interobserver agreement of standard joint count and to compare CE with GSUS findings in patients with early RA* *p. 54*
Population sample size	44
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Early RA (< 2 years)
Population baseline characteristics	72.7% women Mean (SD) age 53 (9.8) years Mean (SD) disease duration 17 (3.8) months
Population treatment (e.g. bDMARDS or cDMARDs)	Prior cDMARDs in all, bDMARDs in 17%
Joints assessed	Acromioclavicular, glenohumeral, elbow, wrist (radiocarpal), MCP, PIP of hands, knee and ankle (tibiotalar), MTP and hip joints, bilaterally
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS US examinations were performed using an AU5 'Harmonic' (Esaote Biomedica, Genoa, Italy) equipped with two broadband linear probes (7.5–10 and 10–14 MHz). Patients were evaluated using the Naredo *et al.*[54] scanning protocol. Joint inflammation was detected using OMERACT definitions. No details of the scoring system were reported
Who conducted US	A rheumatologist experienced in US
Comparator CE details	TJC, SJC
Who conducted comparator CE	Two rheumatologists
Primary outcome of study	Interobserver agreement of standard joint count
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*CE is far from optimal for assessing joint inflammation in patients with early RA . . . US can considerably improve the detection of signs of joint inflammation in terms of both sensitivity and reliability* *p. 54*

SD, standard deviation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 67 Data extraction table: Saleem and Brown[124]

First author (study name)	Saleem[124]
Year	2011
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	To assess: *the accuracy of more stringent remission criteria for indicating the absence of inflammation, using US imaging assessment of synovitis as the gold standard* *p. 792*
Population sample size	128
Population diagnosis of RA	ACR criteria
Population eligibility details (e.g. early RA, remission)	Clinical remission
Population baseline characteristics	Sex NR Mean (SD) age 54 (14) years Median (1st to 3rd quartile) disease duration 8 (5–13) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	DMARD or combination of TNFi and MTX
Joints assessed	Dominant-hand MCP joints 2–5 and wrist (640 joints)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US examination was performed using the Phillips ATL HDI 3000 machine (DMARD group) and the HDI 5000 machine (combination TNFi group) using 10- to 5-MHz and 15- to 8-MHz/12- to 5-MHz 'hockey-stick' transducers, respectively. GSUS and PDUS were used to assess synovial hypertrophy (SH) and synovial vascularity. The presence and location of any SH and PD hyperaemia were recorded according to standardised OMERACT definitions Individual joints were scored for GS SH and PD using a validated 0–3 semiquantitative method (GS: 0 = no SH, 1 = mild SH, 2 = moderate SH, 3 = severe SH; PD: 0 = normal/minimal vascularity, 1 = mild hyperaemia, 2 = moderate hyperaemia, 3 = marked hyperaemia)
Who conducted US	Experienced ultrasonographers
Comparator CE details	Morning stiffness (minutes), 0–100 VAS for patient global assessment of health and disease activity, TJC, SJC, CRP, HAQ-DI, RAQoL
Who conducted comparator CE	Independent trained metrologist
Primary outcome of study	Relationship between clinical scores of remission, imaging-detected synovitis and quality-of-life outcomes
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	Signs and symptoms of inflammation were reduced when more stringent remission criteria were used, but the percentage of joints with PD activity was not reduced, even in those without signs or symptoms. These data suggest that PD is more sensitive than clinical criteria for accurately detecting low but clinically relevant levels of inflammation

SD, standard deviation.

TABLE 68 Data extraction table: Saleem *et al.*[142]

First author (study name)	Saleem[142]
Year	2012
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To determine the clinical, functional and imaging associations of disease flare in patients with RA in remission and any effect on long-term outcomes* *p. 1316*
Population sample size	93
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	In clinical remission as assessed by their consulting rheumatologist; no flares in the last 6 months
Population baseline characteristics	63 (67.7%) female Mean (95% CI) age 56.6 (53.9 to 59.4) years Mean (95% CI) duration of RA 7.0 (4.5 to 9.5) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	cDMARDs, stable treatment for 6 months, no indication for a change in treatment
Joints assessed	Clinical: joints included in the standard 28-joint count US: dominant-hand MCP joints and wrist (intercarpal, radiocarpal, ulnar carpal and distal radioulnar compartments)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS US examination was performed using the Phillips ATL HDI 3000 machine with a 10- to 5-MHz 'hockey-stick' transducer (no other information presented). GSUS and PDUS were used to assess synovial hypertrophy (SH) and synovial vascularity, respectively. The presence and location of any synovial pathology were recorded according to standardised OMERACT definitions Individual joints were scored for GS SH and PD activity using a validated 0–3 semiquantitative method (GS: 0 = no SH, 1 = mild SH, 2 = moderate SH, 3 = severe SH; PD: 0 = normal/minimal vascularity, 1 = mild hyperaemia, 2 = moderate hyperaemia, 3 = marked hyperaemia)
Who conducted US	A single experienced ultrasonographer who was blinded to all other study findings
Comparator CE details	Duration of morning stiffness (minutes), 0–100 VAS for physician and patient global assessment of health and disease activity, SJC, TJC, CRP, (anti-cyclic citrullinated protein antibody, rheumatoid factor), HAQ-DI
Who conducted comparator CE	Independent trained metrologist
Follow-up duration (if relevant)	12 months
Primary outcome of study	Disease flare (defined as any increase in disease activity requiring a change in therapy)
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*Patients without PD activity at baseline have a low likelihood of flaring* *p. 1321*

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 69 Data extraction table: Scheel *et al.*[53]

First author (study name)	Scheel[53]
Year	2005
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To develop an ultrasonographic synovitis scoring system suitable for evaluation of finger joint inflammation in patients with active RA and to compare semiquantitative ultrasonographic scoring with quantitative ultrasonographic measurements* *p. 733*
Population sample size	46
Population diagnosis of RA	ACR criteria
Population eligibility details (e.g. early RA, remission)	Active RA; nine patients had early RA
Population baseline characteristics	37 females, 9 males Mean (range) age 53 (17–75) years Mean (SD) disease duration 8.5 (8.2) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	All treated with DMARDs, including 20 treated with bDMARDs
Joints assessed	Second to fifth PIP and MCP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS US was performed with a HDI 3500 high-end US system (Advanced Technologies Laboratories, Bothell, WA, USA) using a 10- to 5-MHz 'hockey-stick' linear array transducer. Two criteria for active inflammation were evaluated by US: joint effusion (visualised as a black, anechoic area) and thickening of the synovial membrane (visualised as hypo- or hyperechoic structures within the region affected by effusion) Joint effusion and hypertrophy were scored on a 0–3 semiquantitative scale,[52] modified to include both synovitis and effusion in a combined measure [0 = no anechoic, hypoechoic or hyperechoic structure visible (no effusion/hypertrophy); the larger the anechoic structure or extent of synovial hypertrophy seen on US images, the higher the assigned score (1 = minimal effusion/hypertrophy, 2 = moderate effusion/hypertrophy, 3 = extensive effusion/hypertrophy)]
Who conducted US	Investigator experienced in musculoskeletal US
Comparator CE details	DAS28, pain history, CRP, ESR, SJC, TJC
Who conducted comparator CE	No details reported
Primary outcome of study	Optimal US scoring method from six joint combinations [receiver operating characteristic (ROC) curve analysis]
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*US evaluation of finger joint synovitis can be considerably simplified by focusing on the palmar side and by applying semiquantitative grading instead of quantitative measurements* *p. 733*

SD, standard deviation.

TABLE 70 Data extraction table: Spiegel *et al.*[125]

First author (study name)	Spiegel[125]
Year	1987
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	To evaluate: *the correlation of clinically graded joint tenderness and soft tissue swelling with ultrasonographic findings of the same joints and evaluate responses to treatment with fenbufen, a new NSAID* *p. 1283*
Population sample size	6
Population diagnosis of RA	1958 revision of diagnostic criteria for RA
Population eligibility details (e.g. early RA, remission)	Definite or classic RA of ≥ 1 year in duration
Population baseline characteristics	HR; duration of at least 1 year
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	NSAIDs; no NSAIDs for 2 weeks prior to SJC, TJC and US
Joints assessed	Six joints: both shoulders, wrists and knees (transversely and longitudinally)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS Performed using commercial contact B scanners, with 5.0-MHz and 7.5-MHz transducers and using a high-resolution real-time scanner, with a 7.5-MHz transducer. Degree of soft tissue proliferation and amount of effusion were evaluated by a radiologist and graded on a 0–3 semiquantitative scale (0 = none, 1 = mild, 2 = moderate, 3 = severe). Visualisation of an effusion of < 5 mm in depth was diagnosed as a mild soft tissue change. Moderate changes were those effusions between 5 mm and 10 mm in depth, plus visualisation of echogenic soft tissue structures in the joint. Severe changes were effusions of > 10 mm in depth, plus visualisation of echogenic soft tissue abnormalities in the joint
Who conducted US	A radiologist
Comparator CE details	Degree of tenderness and swelling in each of the six joints was graded clinically on the same 4-point scale used for sonograms
Who conducted comparator CE	A rheumatologist
Primary outcome of study	Correlation of joint swelling, joint tenderness and US effusion
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	The patient's subjective index of tenderness, the physician's assessment of joint swelling on physical examination and the findings on ultrasonography were correlated in this study. *Findings suggest that ultrasonography is useful in that it allows an objective and permanent documentation of the amount of synovial effusion and proliferation present . . .* *p. 1288*

NR, not reported.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 71 Data extraction table: Szkudlarek *et al.*[126]

First author (study name)	Szkudlarek[126]
Year	2004
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To compare US with MRI, conventional radiography and CE in the evaluation of bone destruction and signs of inflammation in the MTP joints of patients with RA* *p. 2103*
Population sample size	40 patients (200 MTP joints) (also included healthy control subjects but data reported separately for RA patients)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA
Population baseline characteristics	Female-to-male ratio 3 : 1 Median (range) age 56 (23–78) years Median (range) disease duration 2 (0–20) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	35 patients being treated with DMARDs
Joints assessed	MCP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS US was performed with a General Electric LOGIQ-500 unit using a 7- to 13-MHz linear array transducer. Joint effusion was defined as the compressible anechoic intracapsular area, synovitis as a hypoechoic synovial thickening (non-compressible hypoechoic intracapsular area) and bone erosions as pathological changes in the bone surface of the area adjacent to the joint GSUS findings were scored on a 0–3 semiquantitative scale for bone destruction and joint effusion,[52] including a 0–3 scale for synovial thickening. Grade 3 synovial thickening was further divided into grades 3 and 4 to encompass the more advanced stages of synovial thickening. Grades 0 and 1 were considered to be physiological and grades 2 and 3 (or 4) were considered to be pathological
Who conducted US	The same rheumatologist, who was trained in the examination of the small joints of the hands and feet
Comparator CE details	Swelling and tenderness
Who conducted comparator CE	A consultant rheumatologist
Primary outcome of study	Agreement between US, MRI, CE and radiology
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	Compared with MRI, US was found to be markedly more sensitive and accurate than CE and conventional radiography for the detection and grading of destructive and inflammatory changes in the MTP joints of patients with RA. Evaluation of these joints by US may be of major clinical importance in RA, considering the early and frequent involvement of the MTP joints

TABLE 72 Data extraction table: Szkudlarek *et al.*[127]

First author (study name)	Szkudlarek[127]
Year	2006
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To investigate whether ultrasonography can provide information on signs of inflammation and destruction in RA finger joints that are not available with conventional radiography and CE and which are comparable to the information provided by MRI* *p. 1*
Population sample size	40 patients (158 second to fifth MCP joints and 140 second to fifth PAP joints) (also included healthy control subjects but data reported separately for RA patients)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active RA [20 with established disease (> 2 years)]
Population baseline characteristics	Female-to-male ratio 4 : 1 Mean (range) age 58 (23–79) years Mean (range) disease duration 5 (0–20) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	15 being treated with DMARDs
Joints assessed	Second to fifth MCP joints and second to fifth PIP joints
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS Ultrasonography was performed using a General Electric LOGIQ 500 unit (General Electric, Solingen, Germany) using a 7- to 13-MHz linear array transducer. Each joint was assessed by quadrant for the presence or absence of bone erosions and the presence or absence of signs of inflammation (joint effusion and synovitis) Bone erosion defined as break in bone cortex in the area adjacent to the joint, visualised in two planes; joint effusion defined as compressible anechoic intracapsular area; and synovitis defined as uncompressible hypoechoic intracapsular area. Changes were scored according to a 0–3 semiquantitative scoring system.[52] Scoring of synovitis was widened to include grade 4 (hypoechoic area bulging out of the joint and stretching over both bone diaphyses of the joint)
Who conducted US	Two radiologists with expertise in musculoskeletal ultrasonography and a rheumatologist with training in the examination of the small joints of the extremities
Comparator CE details	Swelling and tenderness
Who conducted comparator CE	The consultant rheumatologist on duty
Primary outcome of study	Agreement between US, MRI, CE and radiology
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*This study shows that ultrasonography has the potential to improve the assessment of patients with RA* *p. 1*

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 73 Data extraction table: Taniguchi *et al.*[128]

First author (study name)	Taniguchi[128]
Year	2014
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	To examine the usefulness of maximum intensity projection MRI for RA in the hand
Population sample size	30
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA, not taking bDMARDs or oral steroids (18 patients in remission with a DAS28-CRP of < 2.3)
Population baseline characteristics	25 women and 5 men Mean (range) age 61.5 ± 9.5 (38–81) years Mean (range) disease duration 12.5 ± 11.5 years (4 months to 45 years)
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	No bDMARDs or oral steroids, MTX in 27 patients
Joints assessed	US on 60 wrists and 300 MCP joints. (CE of 28 joints of DAS28)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Ultrasonography was performed using a HI VISION Avius (Hitachi Medical Corporation, Tokyo, Japan) with a linear-type (14–6 MHz) probe GS images with low echo regions within joints were considered to indicate synovitis. The PD frequency was set at 7.5 MHz and the pulse repetition frequency was set between 800 Hz and 1000 Hz. For PD images, each joint was scored on a semiquantitative scale (0–3) with a score of grade 1 or higher taken as positive (grade 0 = no flow in the synovium, grade 1 = single-vessel signals, grade 2 = confluent-vessel signals in less than half of the area of the synovium, grade 3 = vessel signals in more than half of the area of the synovium[52])
Who conducted US	Orthopaedic surgeon trained in US examination of the small joints of rheumatoid hands. Two orthopaedic surgeons specialising in RA scored the joints independently
Comparator CE details	DAS28-CRP
Who conducted comparator CE	NR
Primary outcome of study	Sensitivity of CE with reference MRI or PDUS
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	CE: *showed low sensitivity and high specificity compared with both MRI and PDUS images. A statistically significant correlation between the scores of MRI and PDUS images was found* *p. 911*

NR, not reported.

TABLE 74 Data extraction table: Vlad *et al.*[129]

First author (study name)	Vlad[129]
Year	2015
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	To investigate the responsiveness of tenosynovitis of the wrist and hands compared with the responsiveness of synovitis in a 6-month follow-up period, by ultrasonography (US) in a cohort of active RA patients starting biologic therapy; to compare the responsiveness of finger flexor tenosynovitis with the responsiveness of wrist extensor tenosynovitis by US; and to describe the subclinical synovitis and tenosynovitis by US in RA patients in clinical remission
Population sample size	57 (55 at follow-up)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Active disease; history of treatment with two different cDMARDs at a maximal dosage for at least 3 months each (initiating bDMARDs)
Population baseline characteristics	50 female patients Mean (SD, range) age 55.28 (10.13, 26–75) years Mean (SD, range) disease duration 113.9 (105.2, 6–414) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Initiation of biological DMARDs
Joints assessed	Bilateral wrists at the level of the radius, lunate and capitate bones from the dorsal side, MCP joints 2–5 and PIP joints 2–5 from both the dorsal and the volar sides
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Ultrasonography was performed using Esaote MyLab 25, 50 and 70 equipped with a linear transducer with 12- to 18-MHz frequency, adjusted for small joint evaluation. For PD, machines were set for maximal sensitivity to detect blood flow (Doppler frequency 6.5–7.5 MHz, pulse repetition frequency 500–750 Hz, low wall filters). All settings were maintained constant throughout the study Joint synovitis was graded on separate 0–3 semiquantitative scales for GS and PD.[52] Synovitis on GS was defined as the presence of abnormal hypoechoic material within joint recesses. PD synovitis was defined as the presence of Doppler signals inside the intra-articular hypoechoic area. In addition to the semiquantitative scale, a binary grading was performed for joint synovitis [0 = no synovitis, 1 = present synovitis (including grades 1, 2 and 3 on the semiquantitative scale)], for both GS and PD US remission was defined as the absence of GS and PD synovitis in all examined joints. Low level of imaging activity was defined less strictly, allowing gradually one to a maximum of two joints or tendons with positive synovitis or tenosynovitis, binary graded (GS and PD global scores for binary evaluation of synovitis and tenosynovitis ≤ 1 or ≤ 2)
Who conducted US	Rheumatologist with ≥ 5 years' experience in musculoskeletal US (at each site)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 74 Data extraction table: Vlad *et al.*[129] (*continued*)

First author (study name)	Vlad[129]
Comparator CE details	TJC, SJC, patient VAS for pain and general disease evaluation, physician VAS evaluation, ESR, CRP, CDAI and SDAI
Who conducted comparator CE	Rheumatologist. One joint assessor at each of five sites (same one for the duration of the study)
Follow-up duration (if relevant)	6 months
Primary outcome of study	Responsiveness of synovitis (vs. responsiveness of tenosynovitis)
Outcome(s) reported in main body of report	Diagnostic (SRM)
Study authors' conclusions	*Tenosynovitis US scoring in RA may be as good as synovitis scoring for characterisation of disease activity and responsiveness* *p. 352*

SD, standard deviation.

TABLE 75 Data extraction table: Wakefield *et al.*[130]

First author (study name)	Wakefield[130]
Year	2008
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To compare CE and US with high-field MRI for the detection of rearfoot and midtarsal joint synovitis and tenosynovitis of the ankle tendons in patients with established RA* *p. 1678*
Population sample size	22
Population diagnosis of RA	Modified ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA with symptoms of midfoot and rearfoot disease
Population baseline characteristics	14 female, 8 male Mean (range) age 52 (33–70) years Mean (range) disease duration 6.8 (1–20) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	All patients were on stable doses of NSAIDs. Ten were taking MTX monotherapy, one was taking SSZ, one was taking HCQ, one was taking gold and six were on a combination of MTX and a TNFi
Joints assessed	Right tibiotalar, subtalar, talonavicular and calcaneocuboid joints (tendons also examined)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS ATL HDI (Advanced Technologies Laboratories, High Definition Imaging, Bothell, WA, USA) 3000 machine employing a 10- to 5-MHz linear array 'hockey-stick' transducer Synovitis was defined as an abnormal hypoechoic area within the joint, compatible with the OMERACT US group definition
Who conducted US	An experienced sonographer; a second experienced sonographer examined five patients
Comparator CE details	CE for synovitis (and tenosynovitis)

TABLE 75 Data extraction table: Wakefield *et al.*[130] (*continued*)

First author (study name)	Wakefield[130]
Who conducted comparator CE	A podiatrist; a second podiatrist examined five patients
Primary outcome of study	Sensitivity of US and CE for detecting synovitis and tenosynovitis, with a reference standard of MRI
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*CE was sensitive but US was more specific in identifying hindfoot pathology in RA compared with the reference standard of MRI* *p. 1678* The authors suggest: *a need for standardisation of acquisition and interpretation of US images of the hindfoot* *p. 1678*
Outcome data additional to main report	Interobserver variability between ultrasonographers was low, although this was based on five patients only

TABLE 76 Data extraction table: Wakefield *et al.*[143]

First author (study name)	Wakefield[143]
Year	2007
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To examine the longitudinal relationship between clinical remission and imaging remission with the hypothesis that it may be more appropriate to aim for persistent absence of imaging synovitis rather than clinical remission* *p. 1564*
Population sample size	10
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	Patients with early RA (< 12 months of symptoms). None of the patients had ever received DMARDs or biological treatments and none had received corticosteroids during the preceding month
Population baseline characteristics	5 male, 5 female Median (range) age 52.5 (21–78) years Median (range) disease duration 6 (3–11) months
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	None of the patients had ever received DMARDs or biological treatments and none had received corticosteroids during the preceding month. At the start of the study, patients were started on a bDMARD (IFX) and rapidly escalating MTX
Joints assessed	CE: 28 joints of DAS. US: 42 joints (bilateral glenohumeral, elbow, wrist, MCP, PIP, knee, tibiotalar, midtarsal and MTP joints)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 76 Data extraction table: Wakefield *et al.*[143] (*continued*)

First author (study name)	Wakefield[143]
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	GSUS and PDUS Philips HDI 5000 (Phillips Medical Systems, Best, the Netherlands) employing either 12- to 5- or 13- to 7-MHz linear transducers Individual joints were scored for synovitis using a semiquantitative scoring method on a 0–3 scale for both GS and PD (not referenced). Scores were expressed per joint and as a total score. Absence of imaging synovitis was arbitrarily defined as both a total GS and a total PD score of 0
Who conducted US	An experienced sonographer
Comparator CE details	DAS28 assessment (also HAQ and the RAQoL) questionnaire and CRP). Clinical remission was defined as a DAS28 of < 2.6 and clinical response was defined as a decrease in DAS28 of > 1.2
Who conducted comparator CE	Treating clinician
Follow-up duration (if relevant)	46 weeks
Primary outcome of study	Spearman's rho correlation coefficients between baseline GS, PD scores and DAS28, time to clinical remission and time spent in clinical remission
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*Even when clinical remission is achieved with TNFis, the absence of imaging synovitis may not be achieved* *p. 1566*

TABLE 77 Data extraction table: Xiao *et al.*[131]

First author (study name)	Xiao[131]
Year	2014
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To investigate the value of US for diagnosing synovitis associated with RA* *p. 767*
Population sample size	46 (also healthy control subjects, but data reported separately for RA patients)
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA patients
Population baseline characteristics	31 females, 15 males Age [average, no further details (range)] 51 (21–67) years Mean (SD, range) disease duration 8.7 (8.7, 0.08–30.0) years

TABLE 77 Data extraction table: Xiao *et al.*[131] (*continued*)

First author (study name)	Xiao[131]
Population treatment (e.g. bDMARDS or cDMARDs)	NR
Joints assessed	828 joints – CP 2–5, PIP 2–5 and wrist joints bilaterally
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS Ultrasonography was performed using the GE Logiqbook XP colour ultrasonic apparatus (General Electric Company, Fairfield, CT, USA) with high-frequency linear matrix probes, the centre frequency of which was 11 MHz OMERACT diagnostic criteria were used for reference. A 0–3 semiquantitative score[59] was used for PDUS (0 = no blood image; 1 = single-vessel image; 2 = several vessels that are partially confluent; 3 = confluent vessels covering more than half of the area of the synovium)
Who conducted US	Two skilled doctors who had undergone specialist training in US
Comparator CE details	Tenderness and/or swelling
Who conducted comparator CE	Experienced doctors from rheumatology departments
Primary outcome of study	US synovitis detection
Outcome(s) reported in main body of report	Diagnostic
Study authors' conclusions	*US is a valid method for diagnosing early-stage synovitis, with high-accuracy cut-off points for MCP, PIP and wrist joints set at 2.5, 2.6 and 5.2 mm* *p. 767*

NR, not reported; SD, standard deviation.

TABLE 78 Data extraction table: Yoshimi *et al.*[144,145]

First author (study name)	Yoshimi[144,145]
Year	2013 and 2014
Abstract or full paper	Full paper
Study design	Prognostic
Study objective	*To assess whether US can predict progressive joint destruction during clinical remission of RA* *Yoshimi* et al.,[144] *p. 456*
Population sample size	31
Population diagnosis of RA	ACR 1987 criteria
Population eligibility details (e.g. early RA, remission)	RA in clinical remission (clinical remission criteria using DAS28-ESR of < 2.6 or DAS28-CRP of < 2.3) for at least 2 months
Population baseline characteristics	Male 4, female 27 Mean (SD) age 55.2 (13.4) years Median (range) disease duration 5 years 0 months (2 months to 6 years 5 months)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 78 Data extraction table: Yoshimi *et al.*[144,145] (*continued*)

First author (study name)	Yoshimi[144,145]
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	Biological therapy: $n = 13$ (IFX, $n = 4$; ETN, $n = 9$); DMARDs: $n = 28$ (MTX, $n = 23$; SSZ, $n = 6$; tacrolimus, $n = 1$); steroids: $n = 9$ (up to 5 mg/day); drug free: $n = 1$
Joints assessed	28 joints (as for DAS28) for clinical assessment; 22 of these joints (excluding bilateral glenohumeral, elbow and knee joints) for US
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	PDUS and GSUS Aplio SSA-700 A apparatus (Toshiba, Tokyo, Japan) with 12-MHz linear array transducers PD signals were graded on a 0–3 semiquantitative scale [0 = absent (no synovial flow); 1 = mild (single-vessel signal or isolated signals); 2 = moderate (confluent signals in less than half of the synovial area); 3 = marked (signals in more than half of the synovial area)].[54] For the PIP and MCP joints, GS images were scored on a 0–3 semiquantitative scale [0 = none (no synovial thickening); 1 = mild (filling the angle between the periarticular bones without bulging over the line linking the tops of the bones); 2 = moderate (synovial thickening bulging over the line linking the tops of the periarticular bones but without extension along the bone diaphysis); 3 = severe (synovial thickening bulging over the line linking the tops of the periarticular bones and with extension to at least one of the bone diaphyses)].[52] For wrists, GS images were scored on a 0–3 semiquantitative scale (0 = none, 1 = mild, 2 = moderate, 3 = severe) on subjective appraisal
Who conducted US	Experienced rheumatologists
Comparator CE details	TJC, SJC, CRP, ESR, matrix metalloproteinase, DAS28-ESR and DAS28-CRP
Who conducted comparator CE	Rheumatologists
Follow-up duration (if relevant)	2 years
Primary outcome of study	Relationship between US at baseline and radiographic progression at 2 years
Outcome(s) reported in main body of report	Prognostic
Study authors' conclusions	*PDUS detects synovitis causing joint destruction even when the patient is in clinical remission* *Yoshimi* et al.,[144] *p. 456*

Prognostic sensitivity (no clinical comparator sensitivity data reported)

Study	Population	Measure being assessed for prediction	US measure	Sensitivity, %	Specificity, %	PPV, %	NPV, %
Yoshimi 2013[144]	22 in RA clinical remission	Radiographic progression at 2 years	Total PD scores of > 1	100 (95% CI 59.0 to 100)	73.3 (95% CI 44.9 to 92.2)	63.6	100

SD, standard deviation.

TABLE 79 Data extraction table: Zuffery[99,132]

First author (study name)	Zuffery[99,132]
Year	2014
Abstract or full paper	Full paper
Study design	Diagnostic
Study objective	*To evaluate whether RA patients considered to be in remission according to clinical criteria sets still have persisting US synovitis. Also, to evaluate the capacity of the US score to discriminate between the patients with clinically active disease and those in remission* *Zuffery,[132] p. 220* *To evaluate the correlation between clinical measures of disease activity and a US scoring system for synovitis applied by many different ultrasonographers in a daily routine care setting within the Swiss registry for RA (SCQM) and further to determine the sensitivity to change of this US score* *Zuffery,[99] p. 35*
Population sample size	307[99] 536 in a cross-sectional cohort and 183 in a longitudinal cohort[132]
Population diagnosis of RA	NR
Population eligibility details (e.g. early RA, remission)	Clinical remission
Population baseline characteristics	Active disease ($n = 167$): 82% female; mean (SD) age 57 (13) years; median (IQR) disease duration 6.8 (2.4–16.5) years DAS remission ($n = 129$): 70% female; mean (SD) age 55 (15) years; median (IQR) disease duration 6.1 (3.0–12.2) years ACR/EULAR remission ($n = 69$): 67% female; mean (SD) age 53 (15) years; median (IQR) disease duration 4.4 (2.6–10.2) years DAS remission only ($n = 71$): 78% female; mean (SD) age 56 (15) years; median (IQR) disease duration 7.2 (3.0–16.5) years
Population treatment at baseline (e.g. bDMARDS or cDMARDs)	48–59% on biological therapy[99] 57% on biological therapy[132]
Joints assessed	22 joints (knee, elbow, wrist and fingers bilaterally)
Type(s) of US and US details (including the machine used, scoring system used and diagnostic cut-off if applicable)	Multiplanar GSUS and Doppler mode (PDUS) The SONAR score is a semiquantitative score employing both multiplanar GSUS and Doppler mode (PDUS). Synovitis was graded from 0 to 3 according to the OMERACT consensus. PD scoring was graded on a 0–3 scale according to OMERACT recommendations. Some operators did not look for PD on the dorsal aspect of the joints when only grade 1 synovitis on GSUS was present on the volar side. Those missing values were assumed to be equivalent to grade 0 on PD
Who conducted US	30 ultrasonographers trained on the SONAR score
Comparator CE details	ACR/EULAR, DAS28, ESR and CRP
Who conducted comparator CE	Could be performed by the ultrasonographer or by another physician
Follow-up duration (if relevant)	Mean (SD) follow-up duration 11.7 (5.6) months
Primary outcome of study	Correlation between clinical and US data
Outcome(s) reported in main body of report	Diagnostic (SRM)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 79 Data extraction table: Zuffery[99,132] (*continued*)

First author (study name)	Zuffery[99,132]
Study authors' conclusions	*This observational study confirms that many patients considered to be in clinical remission according the DAS and the ACR/EULAR definitions still have residual synovitis on US* *Zuffery,[99] p. 35* *The SONAR score is practicable and . . . demonstrates significant correlations with the degree of as well as change in disease activity as measured by DAS. On the level of the individual, the US score shows many discrepancies and overlapping results exist* *Zuffery,[132] p. 220*

IQR, interquartile range; NR, not reported; SD, standard deviation.

Appendix 6 Quality assessment

Relevant items to assess study bias were taken from Karsh *et al.*,[84] the QUADAS tool,[82] GATE[83] and the QUIPS tool.[81]

These are validated, widely used tools. Other tools are available. The QUADAS-2 tool[280] is a more recent version of the QUADAS tool, which could have been used if the QUADAS tool was not sufficient, namely if there had been concerns about applicability of studies to the review. Applicability had been addressed in the review inclusion/exclusion criteria.

As there were several types of study included in the review, not all items were applicable to all study designs. Note that when studies included healthy control subjects in addition to RA patients, in line with the exclusion criteria, studies were included only if outcome data were reported separately for the RA subgroup. As stated previously, there is currently no gold standard/reference standard for detecting synovitis objectively.

Diagnostic study quality assessment items taken from the QUADAS tool

Many items from the QUADAS tool[82] did not apply as the inclusion/exclusion criteria for the review meant that all studies would meet the following criteria:

- Was the spectrum of patients representative of the patients who will receive the test in practice?
- Is the reference standard likely to correctly classify the target condition?
- Did patients receive the same reference standard regardless of the index test result?
- Was the reference standard independent of the index test?
- Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?

Items considered relevant to this review were:

- Were selection criteria clearly described?
- Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?
- Was the execution of US described in sufficient detail to permit replication of the test?
- Was the execution of the CE described in sufficient detail to permit replication of the test?
- Were the US results interpreted without knowledge of the results of the CE?
- Were the CE results interpreted without knowledge of the results of US?
- Were uninterpretable test results reported?
- Were withdrawals from the study explained?

Prognostic study quality assessment items taken from GATE

- Was the outcome measure assessment blinded?
- Was the follow-up time sufficiently long (to detect important prognostic factors)?

Prognostic study quality assessment items taken from GATE and the QUIPS tool

- Were the prognostic factors clearly defined?
- Was the outcome measure clearly defined?

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Prognostic study quality assessment items taken from the QUIPS tool

- Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders).
- The method and setting of outcome measurement is the same for all study participants.
- Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment).

Quality assessment items taken from Karsh *et al.*[84]

Most quality assessment items from Karsh *et al.*[84] did not apply as they referred to trials of therapy. RA-specific quality assessment items taken from Karsh *et al.*[84] were:

- Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?
- Were joint assessments performed by a trained, independent blinded joint assessor? (Broken down into questions about independence of testing and training for clinical and US examinations, for which we included experienced staff as trained).

Quality assessment forms

When studies were diagnostic studies or prognostic studies, the level of study in the hierarchy according to Merlin *et al.*[159] was recorded. This is as follows:

- diagnostic study hierarchy of evidence –
    - level II – diagnostic test accuracy studies with an independent, blinded comparator of a valid reference standard, tested on consecutive patients
    - level III-1 – comparative studies with an independent, blinded comparator of a valid reference standard tested on non-consecutive patients
    - level III-2 – comparative studies not meeting criteria for higher-level evidence
    - level III-3 – diagnostic case–control studies
- prognostic study hierarchy of evidence
    - level II – prospective cohort study
    - level III-1 – all-or-none study
    - level III-2 – single arm of a RCT
    - level III-3 – retrospective cohort study.

TABLE 80 Quality assessment: Backhaus *et al.*[69]

First author (study name)	Backhaus[69]
Year	2013
Study design	Prognostic prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	U
Were clinical and US joint assessments conducted independently?	U
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52,53,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	P

N, no; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 81 Quality assessment: Balsa *et al.*[105]

First author (study name)	Balsa[105]
Year	2010
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 82 Quality assessment: Beckers[106]

First author (study name)	Beckers[106]
Year	2004
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N (semiquantitative system for PD, not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 83 Quality assessment: Bhamra *et al.*[89]

First author (study name)	Bhamra[89]
Year	2014
Study design	Treatment decision
If diagnostic study, what level according to hierarchy?[159]	NA
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	P
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	U (no details, not referenced)
Was the conduct of the CE clearly described?	N
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	NA
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	NA
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; NA, not applicable; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 84 Quality assessment: Boyesen and Haavardsholm[134]

First author (study name)	Boyesen[134]
Year	2011
Study design	Prognostic, prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (Naredo *et al.*[54])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	N
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	U
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 85 Quality assessment: Brown *et al.*[135] and Ikeda *et al.*[136]

First author (study name)	Brown,[135] Ikeda[136]
Year	2008, 2007
Study design	Prognostic (cohort)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre-2010) ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT; 0–3 semiquantitative scale[57,282,283])
Was the conduct of the CE clearly described?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 86 Quality assessment: Bugatti *et al.*[147] and Scirè *et al.*[137]

First author (study name)	Bugatti,[147] Scirè[137]
Year	2012
Study design	Prognostic (prospective cohort)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (Meenagh *et al.*,[277] Schmidt *et al.*,[284] Naredo *et al.*[56])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y for US, U for low disease activity
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 87 Quality assessment: Cavet *et al.*[138] and Taylor *et al.*[100]

First author (study name)	Cavet,[138] Taylor[100]
Year	2009, 2004
Study design	Prognostic (part of RCT comparing MTX + IFX with MTX alone in aggressive early RA)
If prognostic study, what level according to hierarchy?[159]	III-2 (prognostic), II (treatment)
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N
Was the conduct of the CE clearly described?	N
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	P
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 88 Quality assessment: Ceponis *et al.*[152]

First author (study name)	Ceponis[152]
Year	2014
Study design	Treatment decision and diagnostic comparison
If diagnostic study, what level according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52,53])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 89 Quality assessment: Ciurtin *et al.*[90,158]

First author (study name)	Ciurtin[90,158]
Year	2013, 2012
Study design	Treatment decision
If diagnostic study, what level according to hierarchy?[159]	NA
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	NA
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	P
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[86])
Was the conduct of the CE clearly described?	N
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	NA
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; NA, not applicable; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 90 Quality assessment: Dale *et al.*[153,154]

First author (study name)	Dale[153,154] (TaSER)
Year	2013, 2014
Study design	Treatment (treat-to-target strategy RCT: DAS28 target vs. DAS28 + musculoskeletal US target)
If intervention study, what level according to hierarchy?[159]	II
Were eligibility criteria clearly described?	P
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	Y
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	NA
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	N
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52])
Was the conduct of the CE clearly described?	P
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	N
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	N
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; NA, not applicable; P, partially; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 91 Quality assessment: Dougados *et al.*[139,149] and Cheung *et al.*[148]

First author (study name)	Dougados[139,148,149]
Year	2013, 2014
Study design	Prognostic, prospective cohort; 4 months of biologic therapy, then further follow-up up to 2 years
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[54,55,281,284])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 92 Quality assessment: Ellegaard *et al.*[101]

First author (study name)	Ellegaard[101]
Year	2011
Study design	Treatment (cohort study)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	NA
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y

N, no; NA, not applicable; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 93 Quality assessment: Filippucci *et al.*[107]

First author (study name)	Filippucci[107]
Year	2006
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre-2010) ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scale; not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	Y
Were the CE results interpreted without knowledge of the results of US?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 94 Quality assessment: Gandjbakhch *et al.*[91]

First author (study name)	Gandjbakhch[91]
Year	2008
Study design	Treatment decision
If diagnostic study, what level according to hierarchy?[159]	NA
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	N
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	U (no details reported; not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; NA, not applicable; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 95 Quality assessment: Garrigues *et al.*[108]

First author (study name)	Garrigues[108]
Year	2013
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	N
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 96 Quality assessment: Gartner *et al.*[109]

First author (study name)	Gartner[109]
Year	2013
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	Y
Was diagnosis of RA confirmed by an earlier version (pre-2010) ACR or EULAR classification criteria for RA?	NA
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	Y
Were the CE results interpreted without knowledge of the results of US?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

NA, not applicable; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 97 Quality assessment: Haavardsholm and Ostergaard[110]

First author (study name)	Haavardsholm[110]
Year	2009
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–4 semiquantitative score; not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 98 Quality assessment: Haavardsholm *et al.*[79]

First author (study name)	Haarvardsholm[79] (ARCTIC)
Year	2015
Study design	Treatment (treat-to-target strategy RCT: DAS44 strategy vs. DAS44 + MSUS strategy)
If intervention study, what level according to hierarchy?[159]	II
Were eligibility criteria clearly described?	P
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	Y
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	NA
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	P
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (Hammer *et al.*[111])
Was the conduct of the CE clearly described?	P
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were uninterpretable test results reported?	U
Were withdrawals from the study explained?	N
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	U
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	U

N, no; NA, not applicable; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 99 Quality assessment: Hammer and Kvien[68] and Hammer *et al.*[111]

First author (study name)	Hammer[68,111]
Year	2010, 2011
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 100 Quality assessment: Hayashi *et al.*[92]

First author (study name)	Hayashi[92]
Year	2014
Study design	Diagnostic study
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	N
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre-2010) ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	U
Were clinical and US joint assessments conducted independently?	U
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	P
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (semiquantitative scoring system; not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	U
Were the CE results interpreted without knowledge of the results of US?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 101 Quality assessment: Horikoshi *et al.*[112]

First author (study name)	Horikoshi[112]
Year	2010
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	P
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	N
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 102 Quality assessment: Ikeda *et al.*[113,146]

First author (study name)	Ikeda[113,146]
Year	2013, 2012
Study design	Prognostic, prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (Naredo *et al.*[56,140])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	U
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 103 Quality assessment: Inanc *et al.*[93]

First author (study name)	Inanc[93]
Year	2014
Study design	Prospective cohort study (prediction of response to bDMARDs by baseline US and clinical features)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	P
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

P, partially; U, unclear; Y, yes.

Note

'Y' indicates higher quality (lower risk of bias).

TABLE 104 Quality assessment: Iwamoto *et al.*[155]

First author (study name)	Iwamoto[155]
Year	2014
Study design	Prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	Y
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	NA
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[140,285])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	Y
Were the CE results interpreted without knowledge of the results of US?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

NA, not applicable; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 105 Quality assessment: Kamishima *et al.*[114]

First author (study name)	Kamishima[114]
Year	2011
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (Meenagh *et al.*[277])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 106 Quality assessment: Kane *et al.*[102]

First author (study name)	Kane[102]
Year	2003
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 107 Quality assessment: Kelly *et al.*[94]

First author (study name)	Kelly[94]
Year	2013
Study design	Treatment decision
If prognostic study, what level according to hierarchy?[159]	NA
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	N
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	U
Was the conduct of the CE clearly described?	N
Were uninterpretable test results reported?	U
Were withdrawals from the study explained?	Y

N, no; NA, not applicable; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 108 Quality assessment: Luengroongroj *et al.*[95]

First author (study name)	Luengroongroj[95]
Year	2015
Study design	Treatment – cohort (treatment tapering)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	N
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Was the conduct of the US examination clearly described?	N
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	U
Was the conduct of the CE clearly described?	N
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 109 Quality assessment: Luukkainen and Saltyshev[115]

First author (study name)	Luukkainen[115]
Year	2003
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	P
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	U
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	P
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	NA
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	NA
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	NA
Was the follow-up time sufficiently long (to detect important prognostic factors)?	NA
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	NA
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	NA

N, no; NA, not applicable; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 110 Quality assessment: Luukkainen and Sanila[103]

First author (study name)	Luukkainen[103]
Year	2005
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 111 Quality assessment: Luukkainen and Sanila[104]

First author (study name)	Luukkainen[104]
Year	2007
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 112 Quality assessment: Mamoto *et al.*[96]

First author (study name)	Mamoto[96]
Year	2013
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	N
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	P
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	N
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scale; not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	U
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 113 Quality assessment: Mandl *et al.*[116,117]

First author (study name)	Mandl[116,117]
Year	2012, 2013
Study design	Ancillary study to RCT
If diagnostic study, what level according to hierarchy?[159]	III-1 (diagnostic), II (intervention study)
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	N (from a trial sample)
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 114 Quality assessment: Naredo *et al.*[55]

First author (study name)	Naredo[55]
Year	2007
Study design	Prognostic, prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[54])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	U

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 115 Quality assessment: Naredo *et al.*[140]

First author (study name)	Naredo[140]
Year	2008
Study design	Prognostic (prospective cohort)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[53,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	U
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 116 Quality assessment: Naredo *et al.*[118]

First author (study name)	Naredo[118]
Year	2013
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[140,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 117 Quality assessment: Naredo *et al.*[156,157]

First author (study name)	Naredo[156,157]
Year	2015, 2014
Study design	Prospective cohort study of treatment prediction, 12 month follow-up, bDMARD tapering at baseline
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; Y, yes.

Note

'Y' indicates higher quality (lower risk of bias).

TABLE 118 Quality assessment: Osipyants *et al.*[97,150]

First author (study name)	Osipyants[97,150]
Year	2013
Study design	Prognostic, prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	U

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 119 Quality assessment: Pereira *et al.*[119]

First author (study name)	Pereira[119]
Year	2015
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52,54,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	Y
Were the CE results interpreted without knowledge of the results of US?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 120 Quality assessment: Ramirez García *et al.*[98]

First author (study name)	Ramirez García[98]
Year	2014
Study design	Prognostic, prospective cohort study
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scoring; not referenced)
Was the conduct of the CE clearly described?	P
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	P
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	U

P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 121 Quality assessment: Reynolds *et al.*[141] and Rees and Pilcher[151]

First author (study name)	Reynolds,[141] Rees[151]
Year	2009, 2007
Study design	Prognostic (prospective cohort)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scoring scales[286])
Was the conduct of the CE clearly described?	P
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	U
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	U

N, no; P, partially; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 122 Quality assessment: Ribbens *et al.*[120]

First author (study name)	Ribbens[120]
Year	2003
Study design	Diagnostic and response to treatment (before-and-after study)
What level of study according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	N (0–3 semiquantitative scale[58,283])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 123 Quality assessment: Riente *et al.*[121]

First author (study name)	Riente[121]
Year	2010
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	Y
Were the CE results interpreted without knowledge of the results of US?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 124 Quality assessment: Riente *et al.*[122]

First author (study name)	Riente[122]
Year	2011
Study design	Diagnostic, blinded comparison among non-consecutive RA patients
If diagnostic study, what level according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	Y
Were the CE results interpreted without knowledge of the results of US?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 125 Quality assessment: Salaffi *et al.*[160]

First author (study name)	Salaffi[160]
Year	2008
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	Y
Were the CE results interpreted without knowledge of the results of US?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 126 Quality assessment: Saleem and Brown[124]

First author (study name)	Saleem[124]
Year	2011
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	Y
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	NA
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[57,75,281,282])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

NA, not applicable; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 127 Quality assessment: Saleem *et al.*[142]

First author (study name)	Saleem[142]
Year	2012
Study design	Prognostic (prospective cohort)
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	U
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 128 Quality assessment: Scheel *et al.*[53]

First author (study name)	Scheel[53]
Year	2005
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 129 Quality assessment: Spiegel *et al.*[125]

First author (study name)	Spiegel[125]
Year	1987
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-1
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	N (1958 criteria – no mention of ACR or EULAR)
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scale; not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 130 Quality assessment: Szkudlarek *et al.*[126]

First author (study name)	Szkudlarek[126]
Year	2004
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	U
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 131 Quality assessment: Szkudlarek *et al.*[127]

First author (study name)	Szkudlarek[127]
Year	2006
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	U
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	U
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 132 Quality assessment: Taniguchi *et al.*[128]

First author (study name)	Taniguchi[128]
Year	2014
Study design	Diagnostic comparison, blinded, not consecutive patients
If diagnostic study, what level according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scale[52])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 133 Quality assessment: Vlad *et al.*[129]

First author (study name)	Vlad[129]
Year	2015
Study design	Diagnostic (responsiveness)
If diagnostic study, what level according to hierarchy?[159]	III-1
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[52])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 134 Quality assessment: Wakefield *et al.*[130]

First author (study name)	Wakefield[130]
Year	2008
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 135 Quality assessment: Wakefield *et al.*[143]

First author (study name)	Wakefield[143]
Year	2007
Study design	Prognostic, prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	Y
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scoring; not referenced)
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	NA
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	Y

N, no; NA, not applicable; Y, yes.

Note

'Y' indicates higher quality (lower risk of bias).

TABLE 136 Quality assessment: Xiao *et al.*[131]

First author (study name)	Xiao[131]
Year	2014
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[59,281])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE?	U
Were the CE results interpreted without knowledge of the results of US?	U
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 137 Quality assessment: Yoshimi *et al.*[144,145]

First author (study name)	Yoshimi[144,145]
Year	2014, 2013
Study design	Prognostic, prospective cohort
If prognostic study, what level according to hierarchy?[159]	II
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	N
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	Y
Were clinical joint assessments conducted by a trained assessor?	Y
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	Y
Was the conduct of the US examination clearly described?	Y
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (0–3 semiquantitative scoring[52,54])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	Y
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	Y
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y
Were the prognostic factors clearly defined (i.e. the variables being assessed for influence on prognosis or therapy response)?	Y
Was the outcome measure clearly defined (i.e. the clinical prognosis or therapy response measure)?	Y
Was the outcome measure assessment blinded (i.e. the clinical prognosis or therapy response measure)?	Y
Was the follow-up time sufficiently long (to detect important prognostic factors)?	Y
Loss to follow-up is not associated with key characteristics (prognostic factor and potential confounders)	Y
Important potential confounders are accounted for in the study design (e.g. matching for key variables, stratification or initial assembly of comparable groups) or in statistical analyses (i.e. appropriate adjustment)	P

N, no; P, partially; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

TABLE 138 Quality assessment: Zuffery[99,132]

First author (study name)	Zuffery[99,132]
Year	2014
Study design	Diagnostic
If diagnostic study, what level according to hierarchy?[159]	III-2
Were selection criteria clearly described?	Y
Was diagnosis of RA confirmed by the 2010 ACR and EULAR classification criteria for RA?	U
Was diagnosis of RA confirmed by an earlier version (pre 2010) of the ACR or EULAR classification criteria for RA?	U
Were clinical joint assessments conducted by a trained assessor?	U
Were US joint assessments conducted by a trained assessor?	Y
Were clinical and US joint assessments conducted independently?	N
Did population recruitment consist of consecutive patients meeting the eligibility criteria?	U
Is the time period between US and CE short enough to be reasonably sure that the synovitis did not change between the two tests?	Y
Was the conduct of the US examination clearly described?	U
Were US results interpreted according to an established scoring system (e.g. OMERACT)?	Y (OMERACT[64,85])
Was the conduct of the CE clearly described?	Y
Were the US results interpreted without knowledge of the results of the CE (i.e. blinded)?	N
Were the CE results interpreted without knowledge of the results of US (i.e. blinded)?	N
Were uninterpretable test results reported?	Y
Were withdrawals from the study explained?	Y

N, no; U, unclear; Y, yes.
Note
'Y' indicates higher quality (lower risk of bias).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 7 Detection of synovitis

This review concentrated on US as the intervention. There is no conclusive gold standard/reference standard for assessing synovitis. This makes interpretation of detection rates difficult. Other imaging techniques are available for detecting synovitis in RA patients. Imaging techniques used in RA include US, conventional radiography, MRI and CT scans. Other imaging techniques may be useful for detecting disease indications other than synovitis, such as bone oedema or erosion.[9,28] It is noted that MRI has a role in diagnosing RA and for detecting other indications such as bone erosion within RA.

Ultrasound is more likely to be practical in assessing synovitis than other imaging techniques. US is more comfortable for the patient than MRI, is less expensive and has an advantage over other imaging techniques, as it can be used immediately after CE to assess symptomatic areas.[30]

As for any imaging technique, the value of detecting subclinical synovitis is determined by the influence of non-clinically detected synovitis on the course of the disease. This review investigated the association of US-detected synovitis with later outcomes, in prognostic studies (see *Chapter 3*). This can be considered as validating the test results.[287] As this review concentrated on US as the intervention, the decision was taken not to compare US-detected synovitis with synovitis detected by other imaging techniques.

In terms of synovitis, a review published in 2013[28] found that scintigraphy and PET detected similar rates of inflammation to CE and that US and MRI found higher rates of inflammation than CE but with similar rates to each other (compared with CE US detected 2.18-fold synovitis and MRI 2.20-fold synovitis). Studies included in the 2013 review overlapped with studies included in this review; however, two studies[128,131] included in this review that were published since the 2013 review reported synovitis detection rates in both US and MRI. Taniguchi *et al.*[128] reported that synovitis was detected by PDUS in 31 out of 60 (51.7%) wrist joints and 23 out of 300 (7.7%) MCP joints; more joints were reported as positive for synovitis when considering maximum intensity projection MRI grades 1 and 2 [47/60 (78.3%) wrist joints and 84/300 (28.0%) MCP joints]. Xiao *et al.*[131] examined 180 joints with MRI and PDUS; the results were positive for synovitis in 86 joints assessed by MRI (47.7%) and 81 joints assessed by PDUS (45.0%). Imaging techniques other than US are not considered further here.

When diagnostic accuracy data were reported, these were usually presented as the diagnostic accuracy of CE, with US as the reference standard. Because of this, *Table 139* shows the sensitivity and specificity of CE with reference US. When TPs, FPs, TNs and FNs were reported or calculable, these are presented in the data extraction tables in *Appendix 5*, as are diagnostic accuracy data for US with reference CE.

Twenty studies[92,96,102–105,107,109,115,118–122,124–128,160] reported sensitivity data. Two of these studies[105,118] used DAS as the clinical comparator, whereas the others used CE for synovitis.

There was seemingly a very wide range of sensitivity (13–95%) and specificity (30–100%) of CE with US as the reference standard across the 20 studies. However, for most studies, CE had a high specificity and low sensitivity when using US as the reference standard. This indicates some agreement between CE and US, with US detecting synovitis in some joints in which CE did not and only a few cases in which CE detected synovitis and US did not. This agrees with the higher detected rates of synovitis for US over CE reported in the majority of studies (*Table 140*).

Sensitivity of CE of swollen and/or tender joints with a reference standard of GSUS or PDUS in 14 studies[92,96,102–104,107,109,115,119,120,124,126–128,133] ranged from 13% to 69% (MCP and PIP joints for Ribbens *et al.*[120]). Most of these studies assessed synovitis in hand and wrist joints, but they also included ankle, elbow and shoulder joints (see *Table 139*). Sensitivity would range from 37% to 69% if three studies with sensitivities at the lower end of the range[84,103,124] were excluded. Studies with sensitivities at the lower end of the range were those in which all patients were in remission (25%;[84] 13% or 20% depending on PDUS

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 139 Clinical examination diagnostic accuracy with US as reference standard

Study	Population[a]	CE assessment of joints	Diagnostic accuracy comparison (CE with reference US)	Sensitivity of CE (95% CI), %	Specificity of CE (95% CI), %
Balsa 2010[105]	97 patients	SDAI of > 5	PDUS	66 (52 to 77)	55 (40 to 69)
	42 joints (PIP, MCP, wrist, elbow, glenohumeral, knee, ankle and midtarsal and MTP joints) (4074 joints in total)	SDAI of > 3.3	PDUS	57 (44 to 70)	74 (59 to 85)
Filippucci 2006[107]	24 patients	Swollen	PDUS[b]	41 (NR)	71 (NR)
	48 wrists	Tender	PDUS[b]	39 (NR)	78 (NR)
	192 examinations				
Gartner 2013[109,133]	90 patients (60 in clinical remission, CDAI of ≤ 2.8), 1320 joints (MCP, PIP and wrist joints)	Swollen	PDUS grade 3[c] and GSUS grade 3[c]	25 (NR)	100 (NR)
Hayashi 2014[92]	208 patients	Swollen and/or tender	GSUS score of ≥ 1[c] or PDUS score of ≥ 2[c]	Wrist 53 (NR), MCP 50 (NR), interphalangeal/PIP 51 (NR)	Wrist 89 (NR), MCP 95 (NR), interphalangeal/PIP 94 (NR)
	416 wrist joints, 2080 MCP joints, 2080 interphalangeal/PIP joints				
Kane 2003[102]	22 patients	Effusion	GSUS effusion[d]	59 (NR)	65 (NR)
	44 knees				
Luukkainen 2003[115]	30 patients	Swollen	GSUS[d]	40 (NR)	77 (NR)
	288 MTP joints				
	30 patients	Swollen	GSUS[d]	46 (NR)	60 (NR)
	60 talocrural joints				
Luukkainen 2005[103]	50 patients	Swollen	GSUS[e]	41 (NR)	92 (NR)
	100 humeroradial joints				
	50 patients	Swollen	GSUS[e]	21 (NR)	99 (NR)
	100 olecranon fossa joints				
Luukkainen 2007[104]	50 patients	Swollen	GSUS[e]	37 (NR)	82 (NR)
	100 glenohumeral joints				
Mamoto 2013[96]	124 patients	Swollen (assessed by patient or physician)	GSUS of ≥ 2[c]	Physician 47 (NR), patient 34 (NR)	NR (NR)
	2728 joints (wrist, MCP and PIP joints bilaterally)		GSUS of ≥ 2 and PDUS of ≥ 1[c]	Physician 52 (NR), patient 37 (NR)	
Naredo 2013[118]	67 patients in clinical remission[f]	DAS28 of > 2.6	GSUS[b]	42 (NR)	100 (NR)
	28 joints (of DAS28)	SDAI of > 3.3	GSUS[b]	71 (NR)	80 (NR)
		DAS28 of > 2.6	PDUS[b]	46 (NR)	69 (NR)
		SDAI of > 3.3	PDUS[b]	77 (NR)	43 (NR)

TABLE 139 Clinical examination diagnostic accuracy with US as reference standard (*continued*)

Study	Population[a]	CE assessment of joints	Diagnostic accuracy comparison (CE with reference US)	Sensitivity of CE (95% CI), %	Specificity of CE (95% CI), %
Pereira 2015[119]	Painless group: 38 patients 304 MCP joints	Swollen	GSUS[b] PDUS[b]	64 (NR) 75 (NR)	55 (NR) 50 (NR)
	Painful group: 34 patients 272 MCP joints	Swollen	GSUS[b] PDUS[b]	66 (NR) 79 (NR)	51 (NR) 47 (NR)
Ribbens 2003[120]	11 patients 20 wrist joints, 110 MCP joints, 103 PIP joints	Swollen	PDUS[b]	Wrist 87 (NR), MCP 64 (NR), PIP 56 (NR)	Wrist 60 (NR), MCP 53 (NR), PIP 76 (NR)
Riente 2010[121]	100 patients 200 knee joints	Swollen and/or painful	GSUS and PDUS[c]	80 (NR)	87 (NR)
Riente 2011[122]	100 patients 200 foot joints	Swollen or painful	US effusion, GSUS and PDUS[c]	79 (NR)	54 (NR)
Salaffi 2008[160]	44 patients 440 PIP joints, 440 MCP joints, 440 MTP joints	Swollen	GSUS	PIP 95 (NR), MCP 95 (NR), MTP 74 (NR)	PIP 80 (NR), MCP 66 (NR), MTP 69 (NR)
Saleem 2011[124]	128 patients in remission (i.e. a DAS28 of < 2.6) 640 MCP and wrist joints	Swollen	PDUS > 0[c] PDUS > 1[c]	13 (8 to 22) 20 (1 to 70)	93 (90 to 95) 92 (90 to 94)
Spiegel 1987[125]	6 patients 36 joints (shoulders, wrists and knees)	Swollen (0 vs. 1–3 on a 0–3 scale)	GSUS[b]	90 (NR)	30 (NR)
Szkudlarek 2004[126]	40 patients 200 MTP joints	Swollen or tender	GSUS[g]	48 (NR)	89 (NR)
Szkudlarek 2006[127]	40 patients 158 MCP and 140 PIP joints	Swollen or tender	GSUS[g]	53 (NR)	94 (NR)
Taniguchi 2014[128]	30 patients (18 in remission; i.e. a DAS28 of < 2.3) 60 wrists and 300 MCP joints	Swollen or tender	PDUS reference[b]	Wrist 69 (NR), MCP 57 (NR)	Wrist 89 (NR), MCP 95 (NR)

NR, not reported.
a Active RA unless otherwise stated.
b > 0 on a 0–3 scale.
c Graded on a 0–3 scale.
d Graded on a 0–1 scale.
e > 0 on a 0–1 scale.
f Neither disease flare nor changes in therapy (including corticosteroid and MTX) in the past 6 months.
g > 0 on a 0–4 scale.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 140 Detection rates of synovitis by US and CE

Author	Population[a]	Joints	Type of US	Detection of synovitis on US	Type of CE	Detection of synovitis on CE
Balsa 2010[105]	97 patients	4074 joints (PIP, MCP, wrist, elbow, bilateral glenohumeral, knee, ankle and midtarsal and MTP joints)	GSUS	92/97 patients (95%)	SDAI of > 5	43/97 patients (44%)
			PDUS	41/97 patients (42%)	SDAI of > 3.3	54/97 patients (56%)
Beckers 2004[106]	21 patients	356 joints (knee, wrist, MCP, PIP, ankle and MTP joints)	PDUS[b]	199/356 joints (42%)	Swollen	266/356 joints (75%)
					Tender	282/356 joints (79%)
Filippucci 2006[107]	24 patients	48 wrists (four time points × 48 wrists totalling 192 examinations)	PDUS[b]	147/192 examinations (77%)	Swollen	74/192 examinations (39%)
					Tender	68/192 examinations (36%)
Garrigues 2013[108]	40 patients	1600 joints (shoulder, elbow, wrist, MCP, PIP, tibiotalar and MTP joints)	GSUS of ≥ 1[c]	477/1600 joints (30%)	Swollen	200/1600 joints (12%)
			PDUS of ≥ 1[c]	279/1600 joints (17%)	Tender	317/1600 joints (20%)
			GSUS of or PDUS of ≥ 1[c]	479/1600 joints (30%)		
Gartner 2013[109,133]	90 patients (60 in clinical remission, CDAI of ≤ 2.8)	1980 finger and wrist joints (1320 remission, 660 active RA)	GSUS of 1–3[c]	89/90 patients (99%)	CDAI of > 2.8	30/90 patients (33%)
			GSUS of 2–3[c]	77/90 patients (86%)		
			PDUS of 1–3[c]	80/90 patients (89%)		
			PDUS of 2–3[c]	37/90 patients (41%)		
	60 in clinical remission (CDAI of ≤ 2.8)	1320 finger and wrist joints	GSUS of 1–3[c]	887/1320 joints (67%)	Swollen	15/1320 joints (1%)
			PDUS of 1–3[c]	269/1320 joints (20%)		
	30 with active RA	660 finger and wrist joints	GSUS of 1–3[c]	436/660 joints (66%)	Swollen	97/660 joints (15%)
Hayashi 2014[92]	208 patients	416 wrist joints	GSUS 1–3[c] or PDUS of 2–3[c]	197/416 joints (47%)	Swollen	52/416 joints (13%)
					Tender	20/416 joints (5%)
					Swollen and tender	58/416 joints (14%)

TABLE 140 Detection rates of synovitis by US and CE (*continued*)

Author	Population[a]	Joints	Type of US	Detection of synovitis on US	Type of CE	Detection of synovitis on CE
		2080 MCP joints	GSUS of 1–3[c] or PDUS of 2–3[c]	331/2080 joints (16%)	Swollen	106/2080 joints (5%)
					Tender	65/2080 joints (3%)
					Swollen and tender	74/2080 joints (4%)
		2080 interphalangeal/ PIP joints	GSUS 1–3[c] or PDUS 2–3[c]	107/2080 joints (5%)	Swollen	43/2080 joints (2%)
					Tender	93/2080 joints (4%)
					Swollen and tender	46/2080 joints (2%)
Horikoshi 2010[112]	6 patients	156 joints [interphalangeal, radiocarpal, intercarpal, radioulnar, PIP (2–5) and MCP joints]	GSUS[b]	74/156 (47%)	Swollen	7/132 joints (5%)
			PDUS[b]	10/156 (6%)		
Kane 2003[102]	22 patients	44 knees	GSUS effusion (effusion vs. no effusion)[d]	27/44 (61%)	CE for swelling (fluctuant fluid observed)	22/44 joints (50%)
Luukkainen 2003[115]	30 patients	288 MTP joints	GSUS[d]	73/288 joints (25%)	Swollen	79/288 joints (27%)
	30 patients	60 talocrural joints	GSUS[d]	13/60 joints (22%)	Swollen	25/60 joints (42%)
Luukkainen 2005[103]	50 patients	100 humeroradial joints	GSUS[e]	29/100 joints (29%)	Swollen	18/100 joints (18%)
	50 patients	100 olecranon fossa joints	GSUS[e]	29/100 joints (29%)	Swollen	7/100 joints (7%)
Luukkainen 2007[104]	50 patients	100 glenohumeral joints	GSUS[e]	27/100 joints (27%)	Swollen	23/100 joints (23%)
Naredo 2013[118]	67 patients clinical remission	28 joints per patient (1876 in total)	GSUS[b]	62/67 patients (93%)	DAS28 of > 2.6	41/67 patients (61%)
			PDUS[b]	35/67 patients (52%)	SDAI of > 3.3	22/67 patients (33%)
Pereira 2015[119]	Painless group: 38 patients	304 MCP joints	GSUS[b]	193/304 joints (63%)	Swollen	174/304 (57%)
			PDUS[b]	88/304 joints (29%)		
	Painful group: 34 patients	272 MCP joints	GSUS[b]	165/272 joints (61%)	Swollen	161/272 (59%)
			PDUS[b]	67/272 joints (25%)		

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 140 Detection rates of synovitis by US and CE (*continued*)

Author	Population[a]	Joints	Type of US	Detection of synovitis on US	Type of CE	Detection of synovitis on CE
Ribbens 2003[120]	11 patients	20 wrist joints	PDUS[b]	15/20 joints (75%)	Swollen	15/20 joints
		110 MCP joints	PDUS[b]	74/110 joints (67%)	Swollen	64/110 joints
		103 PIP joints	PDUS[b]	45/103 joints (44%)	Swollen	39/103 joints
Riente 2010[121]	100 patients	200 knees	GSUS	140/200 joints (70%)	Swollen and/or tender	116/200 joints
			PDUS[c]	115/140 joints (82%)[f]		
Riente 2011[122]	100 patients	200 foot joints	GSUS and PDUS[c]	135/200 joints (68%)	Swollen or painful	137/200 joints
Salaffi 2006[123]	44 patients	440 PIP joints	GSUS[b]	104/440 joints (23.6%)	Swollen	165/440 joints (37.5%)
		440 MCP joints	GSUS[b]	152/440 joints (34.5%)	Swollen	241/440 joints (54.8%)
		440 MTP joints	GSUS[b]	120/440 joints (27.3%)	Swollen	189/440 joints (43%)
Saleem 2011[124]	128 patients in remission (i.e. a DAS28 of < 2.6)	640 MCP and wrist joints	GSUS = 0 and PDUS = 0[c]	22/128 patients (17%)	Swollen	40/128 patients (31%)
			GSUS > 1 and PDUS > 1[c]	72/128 patients (56%)	Tender	23/128 patients (18%)
			PDUS = 0[c]	63/128 patients (49%)	CRP (≥ 5 mg/dl)	46/128 patients (36%)
			PDUS > 1[c]	101/128 patients (79%)		
Scheel 2005[53]	46 patients	184 MCP and 184 PIP joints	GSUS 1–3[c]	86% across MCP and PIP	Swollen	138/368 (37.5%) MCP and PIP joints
						73/184 MCP joints (39.7%)
						65/184 PIP joints (35.3%)
					Tender	137/368 (37.2%) MCP and PIP joints
						64/184 MCP joints (34.8%)
						73/184 PIP joints (39.7%)
Spiegel 1987[125]	6 patients	36 joints (shoulders, wrists and knees)	GSUS[b]	71/101 (70%)	Swollen	85/101 (84%)
Szkudlarek 2004[126]	40 patients	200 MTP joints	GSUS[g]	129/200 joints (65%)	Swollen or tender	70/200 joints (35%)

TABLE 140 Detection rates of synovitis by US and CE (*continued*)

Author	Population[a]	Joints	Type of US	Detection of synovitis on US	Type of CE	Detection of synovitis on CE
Szkudlarek 2006[127]	40 patients	158 MCP and 140 PIP joints	GSUS[g]	194/480 joints (40%)	Swollen or tender	121/480 joints (25%)
Taniguchi 2014[128]	30 patients (18 in remission; i.e. a DAS28 of < 2.3)	60 wrists	PDUS[b]	31/60 joints (52%)	Swollen or tender	25/60 joints (42%)
		300 MCP joints	PDUS[b]	23/300 joints (8%)	Swollen or tender	26/300 joints (9%)
Wakefield 2008[130]	22 patients	22 TTJ	GSUS[d]	10/22 (45%)	Swollen	17/22 (77%)
		22 STJ		Medial aspect 2/22 (9%), lateral aspect 14/22 (64%)		17/22 (77%)
		22 TNJ		12/22 (55%)		17/22 (77%)
		22 CCJ		11/22 (50%)		12/22 (55%)
Xiao 2014[131]	46 patients	368 bilateral MCP joints	PDUS[b]	147/368 joints (40%)	Swollen and/or tender	180/368 joints (49%)
		368 bilateral PIP joints	PDUS[b]	173/368 joints (47%)	Swollen and/or tender	154/368 joints (42%)
		92 wrist joints	PDUS[b]	61/92 joints (66%)	Swollen and/or tender	56/92 joints (61%)

CCJ, calcaneocuboid joint; STJ, subtalar joint; TNJ, talonavicular joint; TTJ, tibiotalar joint.
a Active RA unless otherwise stated.
b > 0 on a 0–3 scale.
c Graded on a 0–3 scale.
d Graded on a 0–1 scale.
e > 0 on a 0–1 scale.
f Only joints in which synovial proliferation was detected were examined using PDUS.
g > 0 on a 0–4 scale.

definition of synovitis[124]) and those investigating elbow joints (21% or 41% depending on joint studied[103]) or shoulder joints (37%[104]) in which US synovitis was scored as a binary variable rather than semiquantitatively (see *Table 3* for scoring systems).

Higher sensitivities (75–95%) were found in six studies[119–122,125,160] (wrist joints for Ribbens *et al.*[120]). Of these, one study, with a sensitivity of 90%, included only six patients and differed from other studies as CE was graded from 0 to 3[125] and one study was primarily investigating the difference between painful and painless joints.[119] In one study, CE had higher sensitivity for wrists ($n = 20$) than for other joints [120] and, in another study, CE had lower sensitivity for MTP joints than for other joints.[160] One study investigated foot joints (sensitivity 79%[122]) and one study investigated knee joints (sensitivity 80%[21]). Detection rates for synovitis were higher for CE than for US (see *Table 140*) in one study of foot joints.[130] This suggests that US is less useful for detecting synovitis in foot joints (not including MTP) than in other joints. However, this was not a within-study comparison and so this cannot be concluded with certainty.

Four of these studies[119,120,122,125] [Ribbens (MCP joints),[120]] also reported lower specificities (30–55%) than the other studies (see *Table 139*), whereas two[121,160] did not (specificities of 66–87%).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The specificity of CE of swollen/tender joints, with US as a reference standard, ranged from 60% to 100% in 14 of the studies, in which joints included were mostly wrists and hand joints but also ankle, knee, elbow and shoulder joints[92,102–104,107,109,115,120,121,124,126–128,133,160] (wrist and PIP joints for Ribbens *et al.*[120]).

Study quality hierarchy did not explain the differences between the studies. Of four studies in which CE had both higher sensitivity and lower specificity than in other trials, only one[119] was of level II hierarchy (blinded comparison among consecutive patients). The other three studies[120,122,125] did not include (or it was unclear if they included) consecutive patients, but were blinded comparisons. However, other studies lower on the hierarchy[92,96,126–128] had results consistent with level II studies.[102–104,107,109,115,118,124,133]

For studies in which US synovitis was scored on a scale from 0 to 1,[103,104,115] sensitivities and specificities of CE were within the range of those reported by studies using a 0–3 or 0–4 scoring system (see *Table 139*). This was also the case for the study using a US score of ≥ 2 as the reference standard[96] (see *Table 139*).

Studies using DAS as the clinical comparator assessed sensitivity and specificity for diagnosis by patient, rather than for each joint (see *Tables 5* and *6*). Like CE for swollen/tender joints, DAS28 had low sensitivity and high specificity when US was the reference standard.[118] SDAI also had high specificity when US was the reference standard; however, SDAI had fairly high sensitivity.[105,118]

When studies reported data separately by joint, sensitivities and specificities for CE were similar for wrist, MCP and PIP joints in the study by Hayashi *et al.*[92] and for MTP and talocrural joints in the study by Luukkainen and Saltyshev[115] Mamoto *et al.*[96] reported that physicians and patients assessed swelling less effectively in MCP than in PIP or wrist joints. CE of olecranon fossa joints had lower sensitivity than CE of humeroradial joints[103] and CE of MTP joints had lower sensitivity than that of PIP and MCP joints,[160] but with a similar specificity. In the study by Ribbens *et al.*,[120] CE had higher sensitivity for wrists ($n = 20$) than MCP ($n = 110$) or PIP ($n = 103$) joints and lower specificity for MCP joints than wrists and PIP joints.

Table 140 shows the types of US and CEs and detection rates in 25 studies.[53,92,102–109,112,115,118–128,130,131] Three studies[105,109,118] (see *Table 6*) investigated DAS. GSUS detected synovitis in more patients than had active disease indicated by SDAI,[105,118] CDAI[109] or DAS28[118] and this was also the case for PDUS, with the exception of PDUS in one study not detecting as many patients as met a SDAI of > 3.3.[105]

Twenty-three studies[53,92,102–104,106–109,112,119–128,130,131] assessed swollen and/or tender joints by CE (see *Table 6*). Nearly all of these studies[53,92,102–104,107–109,112,115,119–122,124,126–128,131] reported a higher rate of detection of synovitis by US than the rate of detection of swelling or tenderness by CE (see *Table 140*). There were mixed results in two studies,[108,119] with higher detection rates with GSUS than CE, but lower detection rates with PDUS than CE; in addition, one study[131] found a lower detection rate for PDUS than CE for MCP joints but a higher detection rate for PDUS than CE for PIP and wrist joints.

Five studies found lower rates of detection by US than by CE, of which two studies investigated ankle and MTP joints,[106,115] one looked at foot joints,[130] one included MTP but also PIP and MCP joints[123] and one was a study in six patients of GSUS of shoulders, wrists and knees.[125] Higher rates of detection by CE than US were suggested to be the result of inflammation from other pathology, such as osteophytes or oedema,[123,130] or external factors such as obesity.[130]

Table 141 shows responsiveness to change of US and CE reported as the SRM. When interpreting their data, Ikeda et al.[113] cited the study by Husted *et al.*,[80] in which a SRM of < 0.2 is considered nil (meaning between –0.2 and + 0.2) and a SRM of > 0.6 is considered relevant, whereas Haavardsholm and Ostergaard[110] cited thresholds introduced by Cohen[288] for effect sizes: < 0.2, trivial; > 0.2 to ≤ 0.5, small; > 0.5 to ≤ 0.8, moderate; > 0.8, large. Using broad definitions (small/moderate/large), most studies reporting the SRM found similar responsiveness for US and CE. GSUS had similar responsiveness to DAS28,[110,111,113,116,132] SDAI,[111,116,117] CDAI,[111,113] SJC,[111] TJC[111] and CRP.[129] PDUS had similar responsiveness to DAS28,[111,113,116,117,132] CDAI,[111,113] SDAI,[111,116117] SJC[111] and CRP.[129] Looking at the types of US within studies, GSUS had similar responsiveness to PDUS.[111,113,114,129,132]

TABLE 141 Ultrasound and CE responsiveness to change

Author	Population[a]	Follow-up	US SRM (95% CI) (p = value)	CE SRM (95% CI)
Haavardsholm 2009[110]	36 patients starting TNFis Wrists	12 months	GSUS –0.37 (–0.89 to 0.16) (NR)	DAS28 –0.36 (–1.26 to 0.35); SDAI –0.59 (–1.31 to –0.02); CDAI –0.55 (–1.17 to –0.04)
Hammer 2010[111]	20 patients starting biologic therapy (ADA) MCP, PIP, elbow, shoulder, hip, knee, ankle and feet joints	12 months	GSUS (78 joints) –1.27 (NR); PDUS (78 joints) –0.89 (NR)	DAS28 –1.32 (NR); CDAI –1.25 (NR); SDAI –1.20 (NR); assessor's global –0.78 (NR); swollen joints (of 40) –1.29 (NR); tender joints (of 40) –1.07 (NR); ESR –0.13 (NR); CRP –0.12 (NR)
Ikeda 2013[113]	66 patients who were taking MTX ($n = 22$), TNFi ($n = 27$) and TCZ ($n = 17$) 28 joints of DAS28	12 weeks	Total GSUS score –1.17 (NR) (MTX –1.19, TNFi –1.48, TCZ –1.05) (NR); total PDUS score –1.37 (NR) (MTX –1.48, TNFi –1.53, TCZ –1.40) (NR)	DAS28-CRP –1.37 (NR) (MTX –1.36, TNFi –1.52, TCZ –1.49) (NR); CDAI –1.20 (NR) (MTX –1.43, TNFi –1.24, TCZ –1.18) (NR)
Kamishima 2011[114]	29 patients, starting TCZ MCP joints	5 months	PDUS (score 0–3) Sum of PDUS grades at 10 MCP joints –0.2595 (NR); PDUS joint index for vascular flow at 10 MCP joints –0.3063 (NR)	CRP –0.6024 (NR); ESR –1.100 (NR); TJC –0.9288 (NR); SJC –0.6506 (NR); DAS28-ESR –1.9692 (NR)
Mandl 2013[116,117]	62 patients ($n = 32$ randomised to ETN + MTX; $n = 30$ randomised to cDMARDs) 28 joints of DAS28	12 weeks	GSUS and PDUS combined measure based on a DAS28 of 0.79 (0.70 to 0.88); GSUS and PDUS combined measure based on SDAI 0.9 (0.52 to 1.17)	Clinical only DAS28 0.87 (NR); clinical only SDAI 1.11 (NR)
Vlad 2015[129]	55 patients starting bDMARDs Wrist, MCP and PIP joints	6 months	Global GSUS synovitis score (score 0–3, global score based on dorsal and volar scores) –1.80 (NR); global PDUS synovitis score –1.30 (NR)	CRP –0.90 (NR)
Zufferey 2014[132]	183 patients (most already on bDMARDs) 22 joints: knees, elbows, wrists and fingers	11.7 months (mean)	GSUS –0.31 (–0.45 to –0.16); log(PDUS + 1) score –0.23 (–0.39 to –0.08)	DAS-CRP –0.50 (–0.65 to –0.35)

NR, not reported.
a Active RA unless otherwise stated.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

One study found that US had higher responsiveness than ESR or CRP.[111] ESR and CRP in this study were the only measures in any study with a SRM of < 0.2 (i.e. counted as a nil effect). Two studies found US to be less responsive than the clinical comparator, with US having lower responsiveness than SDAI or CDAI[110] or than CRP, ESR, TJC, SJC or DAS28-ESR.[114]

Studies suggested that US detects more synovitis than CE alone. Thirty-three studies provided diagnostic data.[53,92,96,102,104,105,107–116,118–122,124–132,160] The majority of these studies reported that US detected more synovitis than CE alone. US could also distinguish synovitis from other pathologies. Most of these studies assessed synovitis in hand and wrist joints, reporting that US detected more synovitis in these joints than CE, and this was also the case for elbow and shoulder joints. Foot and ankle joints were less likely to show an advantage of US over CE. The detection of subclinical synovitis would be useful only if clinically relevant and prognostic studies suggested that US-detected synovitis was associated with radiographic progression.

Appendix 8 Characteristics of included studies

TABLE 142 Characteristics of included studies

Author (study)	Year; abstract or full paper	Data provided to review	Sample size (RA patients recruited)
Haarvardsholm (ARCTIC)	2015; abstract[79]	Treatment	238
Luengroongroj	2015; abstract[95]	Treatment	32
Naredo	2015; full paper[157]	Treatment	77
	2014; abstract[156]		
Bhamra	2014; abstract[89]	Treatment	17
Ceponis	2014; full paper[152]	Treatment	51
Dale (TaSER)	2014; full paper (one study arm)[153]	Treatment	110
	2013; abstract (both study arms)[154]		
Inanc	2014; abstract[93]	Treatment	43
	2014; abstract[93]		
Iwamoto	2014; full paper[155]	Treatment	42
Kelly	2013; abstract[94]	Treatment	109
Ciurtin	2013; abstract[90]	Treatment	39
	2012; abstract[158]		
Ellegaard	2011; full paper[101]	Treatment	109
Gandjbakhch	2008; abstract[91]	Treatment	52
Dougados/Cheung	2014; abstract[148]	Prognostic and treatment	77
	2013; full paper[139]		
	2013; abstract[149]		
Cavet/Taylor	2009; abstract[138]	Prognostic and treatment	24
	2004; full paper[100]		
Ramirez García	2014; abstract[98]	Prognostic	28
Yoshimi	2014; full paper[145]	Prognostic	31
	2013; full paper[144]		
Backhaus	2013; full paper[69]	Prognostic	432
Osipyants	2013; abstract[97]	Prognostic	36
	2013; abstract[150]		
Bugatti/Scirè	2012; full paper[147]	Prognostic	161
	2009; full paper[137]		
Saleem	2012; full paper[142]	Prognostic	93

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 142 Characteristics of included studies (*continued*)

Author (study)	Year; abstract or full paper	Data provided to review	Sample size (RA patients recruited)
Boyesen	2011; full paper[134]	Prognostic	84
Reynolds	2009; full paper[141]	Prognostic	40
	2007; full paper[151]		
Brown/Ikeda	2008; full paper[135]	Prognostic	107
	2007; abstract[136]		
Naredo	2008; full paper[140]	Prognostic	367
Wakefield	2008; full paper[130]	Diagnostic	22
Naredo	2007; full paper[55]	Prognostic	42
Wakefield	2007; full paper[143]	Prognostic	10
Ikeda	2012; abstract[146]	Diagnostic and prognostic	57
	2013; full paper[113]		
Pereira	2015; full paper[119]	Diagnostic	72
Vlad	2015; full paper[129]	Diagnostic	55
Hayashi	2014; abstract[92]	Diagnostic	208
Taniguchi	2014; full paper[128]	Diagnostic	30
Xiao	2014; full paper[131]	Diagnostic	46
Zufferey	2014; full paper[132]	Diagnostic	108
	2014; full paper[99]		
Garrigues	2013; full paper[108]	Diagnostic	40
Gartner	2013; full paper[109]	Diagnostic	90
	2012; abstract[133]		
Mamoto	2013; abstract[96]	Diagnostic	124
Mandl	2013; full paper[116]	Diagnostic	62
	2012; full paper[117]		
Naredo	2013; full paper[118]	Diagnostic	67
Hammer	2011; full paper[68]	Diagnostic	20
	2010; full paper[111]		
Kamishima	2011; full paper[114]	Diagnostic	29
Riente	2011; full paper[122]	Diagnostic	100
Saleem	2011; full paper[124]	Diagnostic	128
Balsa	2010; full paper[105]	Diagnostic	97
Horikoshi	2010; full paper[112]	Diagnostic	6
Riente	2010; full paper[121]	Diagnostic	100
Haavardsholm	2009; full paper[110]	Diagnostic	36
Luukkainen	2007; full paper[104]	Diagnostic	50
Filippucci	2006; full paper[107]	Diagnostic	24

TABLE 142 Characteristics of included studies (*continued*)

Author (study)	Year; abstract or full paper	Data provided to review	Sample size (RA patients recruited)
Salaffi	2006; full paper[123]	Diagnostic	44
Szkudlarek	2006; full paper[127]	Diagnostic	40
Luukkainen	2005; full paper[103]	Diagnostic	50
Scheel	2005; full paper[53]	Diagnostic	46
Beckers	2004; full paper[106]	Diagnostic	21
Szkudlarek	2004; full paper[126]	Diagnostic	40
Kane	2003; full paper[102]	Diagnostic	22
Luukkainen	2003; full paper[115]	Diagnostic	30
Ribbens	2003; full paper[120]	Diagnostic	11
Spiegel	1987; full paper[125]	Diagnostic	6

© Queen's Printer and Controller of HMSO 2018. This work was produced by Simpson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library