Health Technology Assessment 2011; Vol. 15: No. 14
ISSN 1366-5278

Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation

K Malottki, P Barton, A Tsourapas, AO Uthman, Z Liu, K Routh, M Connock, P Jobanputra, D Moore, A Fry-Smith and Y-F Chen

March 2011
10.3310/hta15140

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk

How to obtain copies of this and other HTA programme reports

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

– fax (with **credit card details**)
– post (with **credit card details** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for 'HTA Payment Services' (out-of-hours answer-phone service)

Fax: 0845 812 4001 – put 'HTA Order' on the fax header

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

Paying by credit card
You can order using your credit card by phone, fax or post.

Subscriptions

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of *HTA on DVD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.

Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation

K Malottki, P Barton, A Tsourapas, AO Uthman, Z Liu, K Routh, M Connock, P Jobanputra, D Moore, A Fry-Smith and Y-F Chen*

West Midlands Health Technology Assessment Collaboration, University of Birmingham, Edgbaston, Birmingham, UK

*Corresponding author

Declared competing interests of authors: Anne Fry-Smith was one of the authors of the technology assessment reports compiled to inform the following technology appraisals: TA130 adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis; and TA36 etanercept and infliximab for the treatment of rheumatoid arthritis. She also formed part of the Evidence Review Group for the single technology appraisal of tocilizumab for rheumatoid arthritis in 2009.

Dr David Moore and Kinga Malottki were part of the Evidence Review Group for the single technology appraisal of certolizumab pegol for rheumatoid arthritis.

Dr Pelham Barton constructed the Birmingham Rheumatoid Arthritis Model, which has been used in several National Institute for Health and Clinical Excellence (NICE) technology assessments/appraisals related to rheumatoid arthritis.

Angelos Tsourapas was part of the Evidence Review Group for the single technology appraisal of tocilizumab for rheumatoid arthritis in 2009.

Dr Paresh Jobanputra has participated in several technology assessments/appraisals related to rheumatoid arthritis as the clinical expert of the assessment group. He works as consultant rheumatologist in a large department of rheumatology. His department has taken part in clinical trials of etanercept, adalimumab, rituximab and tocilizumab. The department receives funding for nurses from Schering-Plough Ltd to administer infliximab to patients currently being treated with this drug. Currently he is the chief investigator of a clinical trial comparing adalimumab versus etanercept (ISRCTN95861172). He has, in the past, received support for educational purposes from Wyeth Pharmaceuticals (etanercept), Abbott (adalimumab) and Roche (rituximab). He has also been involved in an advisory board for Roche in relation to tocilizumab and rituximab, and has accepted support

from UCB Pharma for the purposes of study leave to attend an American College for Rheumatology conference in 2009. He has not been involved in any pharmaceutical company submissions to the NICE and has no stocks or shares in any of these companies. Dr Jobanputra was also part of the Evidence Review Group for the single technology appraisal of tocilizumab for rheumatoid arthritis in 2009.

Dr Martin Connock was part of the Evidence Review Group for the single technology appraisal of tocilizumab and certolizumab pegol for rheumatoid arthritis.

Dr Yen-Fu Chen was one of the authors of the technology assessment report that was compiled to inform technology appraisal No. 130 entitled adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis.

Published March 2011
DOI: 10.3310/hta15140

This report should be referenced as follows:

Malottki K, Barton P, Tsourapas A, Uthman AO, Liu Z, Routh K, *et al*. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation. *Health Technol Assess* 2011;**15**(14).

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

NIHR Health Technology Assessment programme

The Health Technology Assessment (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 research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series 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 referees 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.

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 04/26/02. The protocol was agreed in July 2009. The assessment report began editorial review in November 2009 and was accepted for publication in June 2010. 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 referees 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.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

ISSN 1366-5278 (Print)
ISSN 2046-4924 (Online)
ISSN 2046-4932 (DVD)

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.
Printed on acid-free paper in the UK by the Charlesworth Group.

Abstract

Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation

K Malottki, P Barton, A Tsourapas, AO Uthman, Z Liu, K Routh, M Connock, P Jobanputra, D Moore, A Fry-Smith and Y-F Chen*

West Midlands Health Technology Assessment Collaboration, University of Birmingham, Edgbaston, Birmingham, UK

*Corresponding author

Background: Rheumatoid arthritis (RA) is an inflammatory condition that typically causes a symmetrical chronic arthritis. Timely use of disease-modifying antirheumatic drugs (DMARDs) is an essential aspect of disease management, but many patients may not respond even when conventional agents are used optimally.
Objective: To assess the clinical effectiveness and cost-effectiveness of adalimumab (ADA), etanercept (ETN), infliximab (IFX), rituximab (RTX) and abatacept (ABT) when used in patients with RA who have tried conventional agents and have failed to improve after trying a first tumour necrosis factor (TNF) inhibitor.
Data sources: A systematic review of primary studies was undertaken. Databases searched included the Cochrane Library, MEDLINE (Ovid) and EMBASE up to July 2009.
Study selection: Two reviewers assessed titles and abstracts of studies identified by the search strategy, obtained the full text of relevant papers and screened them against inclusion criteria.
Study appraisal: Data from included studies were extracted by one reviewer and checked by a second. The quality of included studies was assessed independently by two reviewers, with any disagreements resolved by discussion and consultation with a third reviewer if necessary.
Results: Thirty-five studies were included in the systematic review: five randomised controlled trials (RCTs), one comparative study, one controlled study and 28 uncontrolled studies. One RCT (REFLEX) demonstrated the effectiveness of RTX. At 6 months significantly more patients treated with RTX achieved American College of Rheumatology (ACR) 20 [relative risk (RR) = 2.85, 95% confidence interval (CI) 2.08 to 3.91] and ACR70 (RR = 12.14, 95% CI 2.96 to 49.86) compared with those treated with the placebo. Differences between groups in favour of RTX were observed at 6 months for mean change from baseline in Disease Activity Score 28 (DAS28) (mean difference –1.50, 95% CI –1.74 to –1.26) and mean change from baseline in Health Assessment Questionnaire (HAQ) score (mean difference –0.30, 95% CI –0.40 to –0.20). One RCT (ATTAIN) demonstrated the effectiveness of ABT. At 6 months significantly more patients treated with ABT achieved ACR20 (RR = 2.56, 95% CI 1.77 to 3.69) and ACR70 (RR = 6.70, 95% CI 1.62 to 27.80) compared with those treated with placebo. Significant differences between groups in favour of ABT were observed at 6 months for mean change from baseline in DAS28 score

(mean difference –1.27, 95% CI –1.62 to –0.93) and mean change from baseline in HAQ score (mean difference –0.34). Twenty-eight uncontrolled studies observed improvement of effectiveness compared with before switching, in patients who switched to ADA, ETN or IFX after discontinued previous TNF inhibitor(s). Four studies were included in the systematic review of cost-effectiveness. Independent economic evaluation undertaken by the assessment group showed that compared with DMARDs, the incremental cost-effectiveness ratios (ICERs) were £34,300 [per quality-adjusted life-year (QALY)] for ADA, £38,800 for ETN, £36,200 for IFX, £21,200 for RTX and £38,600 for ABT. RTX dominates the TNF inhibitors and the ICER for ABT compared with RTX is over £100,000 (per QALY).

Limitations: Paucity of evidence from RCTs for assessing the clinical effectiveness of TNF inhibitors and an absence of head-to-head trials comparing the five technologies.

Conclusions: Evidence from RCTs suggests that RTX and ABT are more effective than supportive care. Data from observational studies suggest that the use of an alternative TNF inhibitor in patients who exhibit an inadequate response to a first TNF inhibitor may offer some benefit, but there remain uncertainties with regard to the magnitude of treatment effects and their cost-effectiveness. Future research should include head-to-head trials comparing the clinical effectiveness and cost-effectiveness of the technologies against each other and emerging biologics.

Funding: This study was funded by the Health Technology Assessment programme of the National Institute for Health Research.

Contents

Glossary ix

List of abbreviations xi

Executive summary xiii

1. Background 1
Description of underlying health problem 1
Current service provision 6
Description of the technologies 7
Degree of diffusion and anticipated costs 10

2. Definition of the decision problem 11
Decision problems 11
Definition of the interventions 11
Population and relevant subgroups 12
Relevant comparators 12
Relevant outcomes 13
Key issues 13
Place of the intervention in the treatment pathway(s) 14
Overall aims and objectives of assessment 14

3. Assessment of clinical effectiveness 15
Methods for reviewing effectiveness 15
Results: quantity and quality of research available 19
Effectiveness of the technologies compared with supportive care 21
Effectiveness of the technologies compared with newly initiated and previously untried conventional disease-modifying antirheumatic drug 86
Effectiveness of the technologies compared with other biologic agents 86
Comparison of effectiveness between technologies (head-to-head comparisons) 86
Subgroup analyses 92
Ongoing studies 112

4. Assessment of cost-effectiveness 115
Systematic review of existing cost-effectiveness evidence 115
Critique of manufacturers' submissions 118
Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions 128
Further critique of manufacturers' models 133
Independent economic assessment 146
Additional sensitivity analysis to assess impact of differences in assumptions between models 161

5. Assessment of factors relevant to the NHS and other parties 167

6. Discussion **169**
Statement of principal findings 169
Strengths and limitations of the assessment 171
Uncertainties 172

7. Conclusions **173**
Implications for service provision 173
Suggested research priorities 173

Acknowledgements **175**

References **177**

Appendix 1 Details of key outcomes used in rheumatoid arthritis trials **193**

Appendix 2 Literature search strategies **197**

Appendix 3 Flow diagram **199**

Appendix 4 Clinical effectiveness: table of excluded studies with rationale **201**

Appendix 5 Cost-effectiveness: table of excluded studies with rationale **207**

Appendix 6 Clinical effectiveness: full paper inclusion/exclusion checklist **209**

Appendix 7 Cost-effectiveness: full paper inclusion/exclusion checklist **211**

Appendix 8 Clinical effectiveness review: data extraction form **213**

Appendix 9 Cost-effectiveness review: data extraction of included studies **219**

Appendix 10 Outcomes not reported in the main text of the report **229**

Appendix 11 Survey of West Midlands rheumatologists **243**

Appendix 12 Withdrawals from treatment with tumour necrosis factor inhibitors **247**

Appendix 13 Scatter plots for comparisons among biologics in the reference case **253**

Appendix 14 Scenario analyses **255**

Appendix 15 Assumptions in the Birmingham Rheumatoid Arthritis Model **271**

Health Technology Assessment **273**

Glossary

American College of Rheumatology (ACR) 20 Defined as a 20% improvement in the counts of the number of tender and swollen joints and at least three items from the following: observer evaluation of overall disease activity; patient evaluation of overall disease activity; patient evaluation of pain; a score of physical disability; and improvements in blood acute phase responses.

ACR50 Defined as a 50% improvement in the counts of the number of tender and swollen joints and at least three items from the following: observer evaluation of overall disease activity; patient evaluation of overall disease activity; patient evaluation of pain; a score of physical disability; and improvements in blood acute phase responses.

ACR70 Defined as a 70% improvement in the counts of the number of tender and swollen joints and at least three items from the following: observer evaluation of overall disease activity; patient evaluation of overall disease activity; patient evaluation of pain; a score of physical disability; and improvements in blood acute phase responses.

Tumour necrosis factor (TNF) inhibitors Biological agents that block TNF activity.

Health Assessment Questionnaire (HAQ) The Health Assessment Questionnaire is designed to assess the physical function of patients. Scores range from 0 (no functional impairment) to 3 (most impaired). Details are provided in *Appendix 1*.

Disease Activity Score (DAS) The DAS is calculated using a formula that includes counts for tender (53 joints) and swollen joints (44 joints), an evaluation by the patient of general health and blood acute phase response. Scores range from 0 (best) to 10 (most active disease).

DAS28 Disease Activity Score 28, similar to DAS above but using only 28 joints for assessment. Scores range from 0 (best) to 10 (most active disease).

List of abbreviations

ABT	abatacept
ACR	American College of Rheumatology
ADA	adalimumab
AE	adverse event
anti-CCP	anti-cyclic citrullinated peptide
ARRIVE	abatacept researched in rheumatoid arthritis patients with an inadequate anti-TNF response to validate effectiveness
ASSURE	abatacept study of safety in use with other rheumatoid arthritis therapies
ATTAIN	abatacept trial in treatment of anti-TNF inadequate responders
AZA	azathioprine
BRAM	Birmingham Rheumatoid Arthritis Model
BSRBR	British Society for Rheumatology Biologics Registry
CEAC	cost-effectiveness acceptability cure
CI	confidence interval
CRP	C-reactive protein
CyA	ciclosporin A
DANBIO	Danish Registry for Biologic Therapies in Rheumatology
DAS	Disease Activity Score
DAS28	Disease Activity Score 28
DMARD	disease-modifying antirheumatic drug
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ESR	erythrocyte sedimentation rate
ETN	etanercept
EULAR	European League Against Rheumatism
Fc	fragment crystallisable
GO-AFTER	GOlimulab After Former anti-tumour necrosis factor Therapy Evaluated in Rheumatoid arthritis
GP	general practitioner
GST	injectable gold
HAQ	Health Assessment Questionnaire
HAQ DI	Health Assessment Questionnaire Disability Index
HCQ	hydroxychloroquine
HR	hazard ratio
HRQoL	health-related quality of life
HUI3	Health Utilities Index Mark 3
IC	indirect comparison
ICER	incremental cost-effectiveness ratio
IFX	infliximab
IgG1	immunoglobulin G1
IR	inadequate response
ITT	intention to treat
i.v.	intravenous
K–M	Kaplan–Meier (curve)
LEF	leflunomide
LTE	long-term extension
MCID	minimal clinically important difference
MS	manufacturer's submission

MTC	mixed-treatment comparison
MTX	methotrexate
NAO	National Audit Office
NICE	National Institute for Health and Clinical Excellence
NOAR	Norfolk Arthritis Register
NSAID	non-steroidal anti-inflammatory drug
OPPOSITE	open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept
Pall	palliation
PCT	primary care trust
PSA	probabilistic sensitivity analysis
QALY	quality-adjusted life-year
QoL	quality of life
RA	rheumatoid arthritis
RCT	randomised controlled trial
RD	risk difference
ReAct	Research in Active Rheumatoid Arthritis
REFLEX	randomised evaluation of long-term efficacy of rituximab in rheumatoid arthritis
RF	rheumatoid factor
RR	relative risk
RTX	rituximab
SD	standard deviation
SF-36	Short Form questionnaire-36 items
SJC	swollen joint count
SSTAG	Southern Swedish Arthritis Treatment Group Registry
STA	single technology appraisal
SUNRISE	study for understanding rituximab safety and efficacy
TB	tuberculosis
TEMPO	Trial of Etanercept and Methotrexate with radiographic Patient Outcomes
TJC	tender joint count
TNF	tumour necrosis factor
TNFα	tumour necrosis factor alpha
TOC	tocilizumab
WTP	willingness to pay

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Note

This monograph is based on the Technology Assessment Report produced for NICE. The full report contained a considerable number of data that were deemed commercial-in-confidence. The full report was used by the Appraisal Committee at NICE in their deliberations. The full report with each piece of commercial-in-confidence data removed and replaced by the statement 'commercial-in-confidence information (or data) removed' is available on the NICE website: www.nice.org.uk.

The present monograph presents as full a version of the report as is possible while retaining readability, but some sections, sentences, tables and figures have been removed. Readers should bear in mind that the discussion, conclusions and implications for practice and research are based on all the data considered in the original full NICE report.

Executive summary

Background

Rheumatoid arthritis (RA) is a common inflammatory condition that typically causes a symmetrical chronic arthritis that causes joint pain, swelling and in some cases a systemic illness. The cause of RA is unknown, but important genetic influences are recognised. The goal of treatment is to achieve remission if patients present with early disease. In later disease, key goals are to control pain and inflammation and thereby reduce functional limitations and the risk of permanent joint damage.

The timely use of disease-modifying antirheumatic drugs (DMARDs) is an essential aspect of contemporary disease management, but many patients may not respond even when conventional agents are used optimally. DMARDs are defined by their ability to modify the disease process such that the risk of progressive joint damage is reduced. Biological agents designed to interrupt the inflammatory pathway have proved to be an important advance in the care of RA patients. The most widely used agents in the UK are tumour necrosis factor inhibitors [adalimumab (ADA, Humira®, Abbott), etanercept (ETN, Enbrel®, Wyeth Pharmaceuticals) and infliximab [(IFX, Remicade®, Schering-Plough Ltd)] and a monoclonal antibody targeting B lymphocytes [rituximab (RTX, Mabthera®, Roche)]. The use of these agents is subject to National Institute for Health and Clinical Excellence (NICE) guidance and all are approved for use provided specific criteria are met. Other agents such as anakinra (an interleukin-1 inhibitor), abatacept [(ABT, Orencia®, Bristol-Myers Squibb Ltd) an antibody that targets cellular interactions], and tocilizumab [(TOC, RoActemra®, Roche) an interleukin-6 inhibitor] are licensed, but currently under assessment or not approved for use by NICE, at the time when this report is being prepared.

This review considers the clinical effectiveness and cost-effectiveness of ADA, ETN, IFX, RTX and ABT when used in patients with RA who have tried conventional agents including methotrexate (MTX) and have failed to improve after trying a first tumour necrosis factor (TNF) inhibitor.

Objectives

The objectives of the assessment report were to assess:

- Whether significant differences in clinical effectiveness and cost-effectiveness exist between ADA, ETN, IFX, RTX and ABT (referred to as 'the interventions' hereafter) when used within their licensed indications in adults with active RA who have had an inadequate response to a first TNF inhibitor prescribed according to current NICE guidance.
- Whether the interventions are clinically effective and cost-effective compared with conventional DMARDs (such as MTX, sulfasalazine and leflunomide).
- Whether the interventions are clinically effective and cost-effective compared with other biologic agents [including TOC, golimumab (Simponi®, Schering-Plough Ltd) and certolizumab pegol (Cimzia®, UCB)].
- Whether the interventions are clinically effective and cost-effective compared with supportive care.

- Whether the clinical effectiveness and cost-effectiveness of the interventions differ significantly between certain subgroups of patients.

Methods

Clinical effectiveness

A systematic review of primary studies (excluding non-randomised studies with less than 20 patients in a treatment arm) of any of the technologies was undertaken. Databases searched included the Cochrane Library, MEDLINE and EMBASE along with other sources from inception up to July 2009. Further data were obtained from dossiers submitted to NICE by the manufacturers of the technologies. Inclusion decisions, quality assessment and data extraction were undertaken according to pre-defined criteria. Owing to heterogeneity between studies and insufficient data, pooling of results was not undertaken.

Cost-effectiveness

A systematic review of published studies on the costs and cost-effectiveness of the technologies for RA patients who had not responded to a TNF inhibitor and a review of the dossiers submitted to NICE by the manufacturers of the technologies were undertaken. In addition, model-based economic evaluations of the cost-effectiveness of the technologies from the perspective of the UK National Health Service (NHS) were carried out.

Results

Clinical effectiveness

Thirty-five studies were included in the systematic review. Five of these were randomised controlled trials (RCTs), one was a comparative study, one was a controlled study and 28 were uncontrolled studies (including one long-term extension of an RCT). Included RCTs compared one of the technologies with placebo and/or ongoing DMARDs/biologics to which the patients have inadequate response. No head-to-head trials directly comparing the five technologies against each other, or comparing the technologies with other biologics or previously untried DMARDs were identified.

Evidence from randomised controlled trials

The effectiveness of RTX was demonstrated in a good-quality RCT (REFLEX). At 6 months, significantly more patients treated with RTX achieved American College of Rheumatology (ACR) 20 [relative risk (RR) = 2.85, 95% confidence interval (CI) 2.08 to 3.91] and ACR70 (RR = 12.14, 95% CI 2.96 to 49.86) compared with those treated with the placebo. Significant differences between groups in favour of RTX were observed at 6 months for mean change from baseline in Disease Activity Score 28 (DAS28) (mean difference –1.50, 95% CI –1.74 to –1.26) and mean change from baseline in Health Assessment Questionnaire (HAQ) score (mean difference –0.30, 95% CI –0.40 to –0.20).

The effectiveness of ABT was demonstrated in a good-quality RCT (ATTAIN). At 6 months, significantly more patients treated with ABT achieved ACR20 (RR = 2.56, 95% CI 1.77 to 3.69) and ACR70 (RR = 6.70, 95% CI 1.62 to 27.80) compared with those treated with the placebo. Significant differences between groups in favour of ABT were observed at 6 months for mean change from baseline in DAS28 score (mean difference –1.27, 95% CI –1.62 to –0.93) and mean change from baseline in HAQ score (mean difference –0.34, insufficient data for calculating 95% CI).

One small RCT (OPPOSITE, $n = 27$) compared switching to IFX versus staying on ETN in patients who had incomplete response to ETN. The study population was not well defined and the comparator was considered inappropriate for this assessment. Two additional RCTs evaluated concurrent use of ABT and TNF inhibitor, which is not recommended in its licence. These studies were not further assessed.

Evidence from observational studies

One non-randomised study found greater but not statistically significant improvement in DAS28 for patients switched to RTX compared with those who switched to an unspecified alternative TNF inhibitor (mean difference –0.35, 95% CI –0.71 to 0.01). Another prospective cohort from the British Society for Rheumatology Biologics Registry showed significantly greater reduction in HAQ score for patients who switched to an unspecified alternative TNF inhibitor compared with switching to non-biologic DMARDs. Twenty-eight uncontrolled studies observed significant improvement in various measures of effectiveness compared with before switching, in patients who switched to ADA, ETN or IFX after discontinued previous TNF inhibitor(s), for various reasons including lack of efficacy, adverse events (AEs) and other reasons.

Subgroup analyses

Evidence from the REFLEX trial suggested that the effectiveness of RTX does not vary significantly depending on reasons for withdrawal, baseline rheumatoid factor status and number of prior TNF inhibitors tried (one vs more than one).

No significant differences in the effectiveness of ABT between subgroups, defined by the number of prior TNF inhibitors (one vs two) and the identity of the prior TNF inhibitor received (ETN vs IFX), were observed in the ATTAIN trial. Some of these subgroup analyses; however, may be underpowered.

Evidence from observational studies showed that the proportion of patients responding to a subsequent TNF inhibitor might vary according to the reason for withdrawal of the previous TNF inhibitor (higher response in patients who withdrew due to intolerance/AEs compared with those who withdrew due to lack of efficacy). The proportion of patients who respond to a subsequent treatment (including TNF inhibitors, RTX and ABT) decreases as the number of prior TNF inhibitor(s) that the patients have tried increases.

Cost-effectiveness

Systematic review

Four studies were included in the systematic review: two studies evaluated ABT and two RTX. One of the RTX studies was UK based. All but one study carried out a cost–utility analysis and reported results in cost per quality-adjusted life-year (QALY). One study carried out a cost-effectiveness analysis and reported results in cost per additional case of low disease-activity state gained (DAS28 less than or equal to 3.2) and cost per additional remission gained (DAS28 less than 2.6). All studies used a decision-analytic model.

Models varied in some important aspects: the type of model used, the sequence of drugs, comparator therapies and time horizon. There was disparity in the selection of perspectives chosen for the analyses. One study reported costs that include both those from a health-care perspective as well as indirect costs and costs of informal care; inclusion of these costs improves the cost-effectiveness of the drug.

A direct comparison of incremental cost-effectiveness ratios (ICERs) between studies was not possible because of the different approaches to modelling, in particular time horizon, country of origin and perspective chosen.

Independent economic assessment

In the reference case all biologic agents were compared with a newly initiated DMARD and against each other. Compared with DMARDs the ICERs were £34,300 (per QALY) for ADA, £38,800 for ETN, £36,200 for IFX, £21,200 for RTX, and £38,600 for ABT. RTX dominates the TNF inhibitors and the ICER for ABT compared with RTX is over £100,000 (per QALY). These results are subject to considerable uncertainty. Important drivers of that uncertainty were found in the scenario analysis to include the assumptions about HAQ progression on biologic treatments, the equation relating HAQ to quality of life, and for comparisons involving RTX the assumed time between treatments. The inclusion of AE costs for biologic therapy made little difference to the results.

Discussion

The limitations predominantly relate to factors outside the control of the assessment group. The major limitation of the assessment was the paucity of evidence from RCTs for assessing the clinical effectiveness of the three TNF inhibitors and a complete absence of genuine head-to-head trials comparing the five technologies against each other, against other biologics or against newly initiated, previously untried DMARDs. Many observational studies were identified. Data from these studies can be confounded by many factors such as patients' baseline disease activity, past history of therapy and methods of selecting and following up patients and analysis of data. Pooling of data was not performed owing to heterogeneity between studies on these respects.

Conclusions

There is lack of good-quality evidence directly comparing the effectiveness of the five technologies against each other. This imposes significant uncertainties with regard to any assessment of their relative cost-effectiveness. Adjusted indirect comparison suggests that there is no significant difference in the effectiveness between RTX and ABT, both of which are supported by good-quality RCT evidence. Existing data do not allow reliable quantification of the effectiveness of TNF inhibitors compared with RTX and ABT. Independent modelling comparing each of the other four technologies with RTX (recommended in the current NICE guidance) suggests RTX dominating ADA, ETN and IFX and an estimated ICER of £131,000 (per QALY) for ABT compared with RTX.

There is lack of evidence comparing the effectiveness of the five technologies to newly initiated, previously untried DMARDs. Independent modelling based on certain assumptions suggests the following ICERs: £34,300 (per QALY) for ADA, £38,800 for ETN, £36,200 for IFX, £21,200 for RTX and £38,600 for ABT.

There is lack of evidence directly comparing the effectiveness of the five technologies with other biologic agents.

Good-quality evidence from RCTs suggests RTX and ABT are more effective than supportive care (including ongoing DMARDs which had provided inadequate control of the disease). Data from observational studies suggest that the use of an alternative TNF inhibitor after patients have experienced an inadequate response to a first TNF inhibitor may offer some benefit, but there remain significant uncertainties with regard to the magnitude of treatment effects and how these translate into cost-effectiveness.

Funding

This study was funded by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1

Background

Summary

Rheumatoid arthritis (RA) is a common inflammatory condition that typically causes a symmetrical chronic arthritis that causes joint pain, swelling and in some cases a systemic illness. The cause of RA is unknown, but important genetic influences are recognised. The goal of treatment is to achieve remission if patients present with early disease. In later disease, key goals are to control pain and inflammation and thereby reduce functional limitations and the risk of permanent joint damage.

Timely use of disease-modifying antirheumatic drugs (DMARDs) is an essential aspect of contemporary disease management, but many patients may not respond even when conventional agents are used optimally. DMARDs are defined by their ability to modify the disease process such that the risk of progressive joint damage is reduced. Biological agents designed to interrupt the inflammatory pathway have proved to be an important advance in the care of RA patients. The most widely used agents in the UK are tumour necrosis factor inhibitors [adalimumab (ADA), etanercept (ETN) and infliximab (IFX)], and a monoclonal antibody targeting B lymphocytes [rituximab (RTX)]. The use of these agents is subject to National Institute for Health and Clinical Excellence (NICE) guidance, and all are approved for use provided specific criteria are met. Other agents such as anakinra (an interleukin-1 inhibitor), abatacept (ABT, an antibody that targets cellular interactions) and tocilizumab (an interleukin-6 inhibitor) are licensed but currently are not approved for use by NICE.

This review considers the clinical effectiveness and cost-effectiveness of ADA, ETN, IFX, RTX and ABT when used in patients with RA who have tried conventional agents including methotrexate and have failed to improve after trying a first tumour necrosis factor inhibitor.

Description of underlying health problem

Clinical features of rheumatoid arthritis

Rheumatoid arthritis (RA) typically begins in middle age and more affects more women than men. Pathologically the disease is characterised by an inflammatory reaction and increased cellularity of the lining layer of synovial joints. Joints such as the proximal interphalangeal joints, metacarpophalangeal joints, wrists, elbows, cervical spinal joints, knees, ankle and foot joints are commonly affected. Affected joints become stiff after periods of inactivity, for example in the morning, become swollen and are variably painful. Other organ systems may also be affected. Patients commonly experience fatigue and blood abnormalities such as anaemia and a raised platelet count. Weight loss, lymph node enlargement, lung diseases (such as pleurisy, pleural fluid and alveolitis), pericarditis, vascular inflammation (vasculitis), skin nodules and eye diseases (reduced tear production or inflammation) may also occur.

The severity of disease, its clinical course and individual responses to treatment vary greatly. Symptoms of RA may develop within days or evolve over many weeks and months.[1] Several distinct patterns of joint disease are recognised, including predominantly small or medium joint disease; predominantly large joint disease; flitting or transient attacks of joint pain (palindromic rheumatism); pain and stiffness of the shoulder and pelvic girdles (polymyalgic disease); disease associated with weight loss and fever (systemic onset); or any combination of these. Pain and disability in early RA is linked to disease severity and to measures of psychological distress.[2]

Disease progression can be relentless or punctuated by partial or complete remissions of variable and unpredictable intervals.

Diagnosis of rheumatoid arthritis

Rheumatoid arthritis is diagnosed from a constellation of clinical, laboratory and radiographic abnormalities. Diagnosis may be obvious or may need specialist assessment or a period of clinical observation. Internationally agreed classification criteria for RA are used widely in contemporary research studies,[3] but it is widely acknowledged that current criteria need to be revised. Current criteria include morning stiffness in joints exceeding 1 hour, physician-observed arthritis of three or more areas, arthritis involving hand joints, symmetrical arthritis, rheumatoid skin nodules, a positive blood test for rheumatoid factor (RF) and radiographic changes typical of rheumatoid disease. Such criteria have limited utility in routine practice and most clinicians diagnose RA without reference to them, as many patients do not meet formal disease classification criteria early in their disease.[4]

Epidemiology

Rheumatoid arthritis affects around 0.5%–1% of the population, three times as many women as men and at age of onset peaks between the ages of 40 years and 70 years. Prevalence of disease at 65 years of age is six times that at 25 years of age. Recent estimates from England and Wales show an annual incidence of 31 per 100,000 women and 13 per 100,000 men, suggesting a decline in recent decades, and a prevalence of 1.2% in women and 0.4% in men.[5] The National Audit Office (NAO) estimates that around 580,000 people have RA in England and that 26,000 patients are diagnosed each year.[6]

Aetiology

A specific cause for RA has not been identified. There appear to be many contributing factors including genetic and environmental influences. Genetic influence is estimated at 50%–60%.[7] The risk of RA in both members of a pair of monozygotic twins is 12%–15% and a family history of RA gives an individual a risk ratio of 1.6 compared with the expected population rate.[8] Many of the genes associated with susceptibility to RA are concerned with immune regulation. For example, the human leucocyte antigen *HLA-DRB1*, which contributes the greatest risk, and *PTPN22*, which makes the second most important genetic contribution in Caucasian populations, are both involved in T-lymphocyte activation and signalling.[9,10]

Infectious agents have been suspected but no consistent relationship with an infective agent has been shown. Sex hormones have also been suspected because of the higher prevalence of RA in women and a tendency for disease to improve in pregnancy. However, a precise relationship has not been identified. A causal link with lifestyle factors such as diet, occupation or smoking has not been shown.

Pathology

Synovial joints occur where the ends of two bones, covered with hyaline cartilage, meet in a region where free movement is desirable. This joint space is encapsulated by a fibrous capsule lined on the inside by a synovial membrane, which functions to secrete fluid to lubricate and nourish hyaline cartilage. In RA the synovial layer of affected joints becomes enlarged as a result of increased cellularity or hyperplasia, infiltration by white blood cells and formation of new blood vessels. This is accompanied by increased fluid in the joint cavity, which contains white blood cells and a high level of protein (an exudate), contributing to the joint swelling. Bony erosions of cartilage and bone occur where synovial tissue meets cartilage and bone. This occurs through the combined actions of synovial tissue (pannus) and resident cartilage and bone cells. Erosions and loss of cartilage are rarely reversible. Such damage, therefore, compromises the structure and function of a normal joint.

Pathogenesis and biological targets in rheumatoid arthritis

A detailed discussion of the pathogenesis of RA is beyond the scope of this report. This subject is reviewed comprehensively elsewhere.[11–13] The synovial membrane in RA contains activated immune cells such as B and T lymphocytes and macrophages. These cells accumulate in synovial tissue. Cells resident in normal joints including synovial fibroblasts, cartilage cells (chondrocytes) and bone cells (osteoclasts) are also activated. Different cytokines, or small proteins, are produced by particular resident and infiltrating cells and aid intercellular communication and influence cellular and tissue behaviour.

A number of cytokines involved in this inflammatory cascade are seen as potential targets for intervention in RA. Drugs that target cytokines and which are licensed or are at a late stage of development currently include anakinra (directed against interleukin-1), tocilizumab [(TOC, RoActemra®, Roche) targeting interleukin-6] and tumour necrosis factor (TNF) inhibitors [including adalimumab (ADA, Humira®, Abbott Laboratories), certolizumab (Cimzia®, UCB), etanercept (ETN, Enbrel®, Wyeth Pharmaceuticals), golimumab (Simponi®, Schering-Plough Ltd) and infliximab (IFX, Remicade®, Schering-Plough Ltd)]. Other agents include abatacept [(ABT, Orencia®, Bristol-Myers Squibb Ltd) also known as CTLA4Ig], which interferes with T-cell activation, and rituximab (RTX, Mabthera®, Roche), which depletes B lymphocytes. Many other potential targets have been identified and a number of novel agents are in clinical trials.[14]

Management of rheumatoid arthritis

The current management of RA is described in detail in a recent National Institute for Health and Clinical Excellence (NICE) guideline.[15] An exhaustive review of management is not provided here. We focus on aspects of disease management that are relevant to the decision problem in this appraisal.

Non-steroidal anti-inflammatory drugs (NSAIDs) and analgesics are commonly used for symptom relief in RA. These drugs do not modify the disease process and key recommendations in NICE guidance centre on minimising the use of NSAIDs because of the potential toxicity of these agents. Corticosteroids are used widely and in a variety of ways. High doses given orally or parenterally (by a variety of routes) are used for the short-term control of disease while waiting for the effects of disease-modifying antirheumatic drugs (DMARDs). Low-dose glucocorticoids are also commonly used either as sole therapy or in combination with DMARDs. Low-dose glucocorticoids have important disease-modifying effects in RA.[16]

Disease-modifying antirheumatic drugs may be divided into conventional DMARDs, which include azathioprine (AZA), ciclosporin A (CyA), gold [GST (given by intra-muscular injection)], hydroxychloroquine (HCQ), leflunomide (LEF), methotrexate (MTX) and sulfasalazine,[17–19] and newer targeted biological agents, described below. Conventional DMARDs such as penicillamine are now used rarely.[18] Conventional DMARDs may be used in combination, especially where there is a poor response to a single DMARD. For example, in early disease MTX is commonly combined with sulfasalazine and HCQ. There are few direct comparisons of individual DMARDs in early disease, but MTX is regarded as the standard against which other drugs should be compared. Most conventional DMARDs have specific dosing and monitoring schedules that require regular visits to a health-care facility and blood tests. How this is managed varies greatly in the UK; for example, in some centres all patients are seen in hospital clinics for drug monitoring whereas in others this occurs largely in the community.

The key objective in early RA management is to achieve remission. Many patients with early inflammatory arthritis (which often does not meet international classification criteria for RA) are able to achieve remission and treatment may be withdrawn in a proportion without relapse.[20] This occurs in randomised trials or therapeutic studies with conventional DMARDs[21–24] used as

monotherapy or in combination, conventional DMARDs combined with TNF inhibitors and also in observational studies. While these reports focus on the excellent outcomes achieved, it is important to recall that 57% of patients with early RA treated with a protocol designed to minimise disease do not achieve remission, around one-third do not achieve their treatment goal and between 31% and 54% of patients have progressive joint damage depending on the treatment strategy after 4 years of treatment.[25]

The NICE RA guidance recommends the use of MTX combined with another DMARD and corticosteroids (used short term) for disease control in early, severe RA. Practice varies; however, and evidence for combining DMARDs is limited and controversial.[26–28] Not all rheumatologists accept the need for DMARD combinations. Some prefer to step up therapy by adding another DMARD to MTX if the disease is inadequately controlled and others choose to replace the first DMARD with a second drug.[29] A necessity for long-term use of multiple medications plainly requires an open dialogue and shared decision making between patients and health professions,[30] especially where expert opinion differs.

In England and Wales patients who have failed to respond to (or tolerate) at least two DMARDs, including MTX at optimal doses, are eligible for TNF inhibitors subject to NICE guidance. Patients who do not respond to TNF inhibitors may be treated with RTX, a monoclonal antibody that depletes B lymphocytes. Other biological therapies such as anakinra, ABT and TOC are not currently approved for use by NICE. The relevant NICE guidance concerned with biologic therapies is described briefly below (see *Current service provision*).

Controlling symptoms of joint pain and stiffness, minimising loss of function, improving quality of life (QoL) and reducing the risk of disability associated with joint damage and deformity are central objectives in the management of RA at all stages. These objectives are not met with drug therapy alone: patients often need advice and support from a multidisciplinary team including specialist nurses, podiatrists, physiotherapists and occupational therapists. Since RA is a heterogeneous disease, which may vary over time, a long-term plan with regular clinical evaluation to assess disease status, disease complications, comorbidity, patient preferences and psychosocial factors is essential and is aided by well-informed and satisfied patients and carers.[31,32] Indeed a key element of a Scottish trial reporting excellent outcomes was frequent specialist review with a focus on tight disease control.[21]

With advanced joint damage surgical intervention such as joint replacement arthroplasty, joint fusion or osteotomy may be necessary. Long-term observations show that around a quarter of patients with RA undergo a total joint arthroplasty.[33] It cannot, of course, be assumed that all such surgery is directly attributable to RA, especially as osteoarthritis is the most prevalent form of arthritis. Other surgical interventions such as removal of synovial tissues and rheumatoid nodules, peripheral nerve decompression (such as in carpal tunnel syndrome) or soft tissue procedures such as tendon release or repair may be necessary at any stage of disease.

Assessment of response to disease-modifying antirheumatic drugs and biologic therapies

ACR response criteria

Modern clinical trials rely on composite end points such as the American College of Rheumatology (ACR) definition of improvement and the Disease Activity Score (DAS). The ACR response requires an improvement in the counts of the number of tender and swollen joints (using designated joints) and at least three items from the following: observer evaluation of overall disease activity; patient evaluation of overall disease activity; patient evaluation of pain; a score of physical disability [such as the Health Assessment Questionnaire (HAQ); see below];

and improvements in blood acute phase responses [e.g. erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP)].

Response is defined as ACR20, ACR50 or ACR70, where figures refer to the percentage improvement of these clinical measures. This creates a dichotomous outcome of responders and non-responders. Achieving an ACR20 response has been regarded as a low hurdle, but in clinical practice patients who achieve this hurdle often gain a worthwhile clinical response, especially in early RA. ACR response criteria are described in more detail in *Appendix 1*.

DAS response criteria

The DAS score is calculated using a formula that includes counts for tender and swollen joints, an evaluation by the patient of general health (on a scale of 0 to 100) and blood acute phase (usually a log of the ESR, but more recently using CRP). DAS response criteria are described in more detail in *Appendix 1*. Originally DAS was based on an assessment of 53 joints for tenderness and 44 joints for swelling. Disease Activity Score 28 (DAS28), based on an evaluation of 28 joints, is used widely in routine clinical practice, partly as a result of NICE guidance on use of TNF inhibitors. DAS28, like DAS, is a continuous scale with a theoretical range from 0 to 10. Thresholds have been suggested for the scale such that a score greater than 5.1 is regarded as indicating high disease activity, a score of less than 3.2 low disease activity and a score of less than 2.6 remission.[34,35] Achieving a DAS28 score of less than or equal to 3.2 and improving the score by greater than 1.2 is regarded to be a good response while achieving a score of less than or equal to 3.2 and improving by greater than 0.6 but less than 1.2 is regarded as a moderate response. Current NICE guidance for TNF inhibitors demands that patients should improve DAS28 by 1.2 in order to justify continuing treatment. It has been suggested that NICE guidance should be altered to allow patients who have attained a moderate response to continue treatment with a TNF inhibitor.[36]

While DAS28 scores are a very valuable tool for assessing treatment responses in groups of patients, scores have important limitations when used for individual patient decisions. For example, DAS28 does not incorporate ankle and foot disease. Thus, a patient with disease localised here may not attain a sufficiently high score to be eligible for a TNF inhibitor. DAS28 also shows poor concordance with clinical judgement (based on a wide range of parameters).[37] In addition, the degree of measurement error in a test–retest reliability study indicates that the faith placed in DAS28 as the sole decision-making tool is misplaced.[38] For example, the smallest detectable difference which should be exceeded if a clinician is to be 95% confident that a change exceeds measurement variability was 1.32 for DAS28.

The Health Assessment Questionnaire

The HAQ is a family of questionnaires designed to assess functional capacity of patients.[39] The most widely used version of HAQ is the modified HAQ (MHAQ) score which comprises eight items such as an ability to dress, get in and out of bed, lift a cup, walk outdoors and wash. MHAQ is reported as an average score across the eight categories such that 0 indicates an ability to achieve tasks without difficulty and 3 reflects an inability to achieve tasks. Scores therefore range between 0 and 3 with an interval of 0.125. Low scores indicate better function. Care is needed in the interpretation of HAQ scores in published studies because there are several modifications to HAQ. The HAQ score is described in more detail in *Appendix 1*.

Radiographic measures

Radiographic outcomes are believed by many to be the most important outcome measure in RA. However, variation in joint inflammation has a more profound and immediate impact on disability compared with the slow and cumulative effect of radiographic damage on disability.[40]

The most commonly used tools for assessing joint damage are the Sharp and Larsen methods and their modifications (see *Appendix 1*), which rely on evaluations of plain radiographs of hands and feet. Plain radiographs are rather insensitive to change but are cheap and widely available. A majority of patients show only mild or no progression on plain radiographs over periods of 1–2 years, highlighting one of their limitations in modern clinical trials.[41]

Prognosis

The impact of RA on an individual can be viewed from a variety of perspectives including employment status, economic costs to the individual or society, QoL, physical disability, life expectancy and medical complications such as extra-articular disease and joint deformity, radiographic damage or the need for surgery. In general, persistent disease activity is associated with poorer outcomes, although in the first 5 years of disease physical function is especially labile. Greater physical disability at presentation is associated with greater disability later in disease. Other factors linked with poorer function include older age at presentation, the presence of rheumatoid nodules, female sex, psychological distress and degree of joint tenderness.[42] Continued employment is related to the type of work and other aspects of the workplace such as pace of work, physical environment, physical function, education and psychological status; work disability is not necessarily linked to measures of disease activity.[43,44] Radiographic damage in RA joints is also influenced by RF status, age, disease duration and extent of disease and perhaps genetic factors.

Life expectancy in RA is reduced and is related to age, disability, disease severity, comorbidity and RF status, in particular.[45–48] For example, a 50-year-old woman with RA is expected to live for 4 years less than a 50-year-old woman without RA.[49] Patients with RA have a significantly increased risk of ischaemic heart disease. Heart disease is the principal reason for an approximately 60% increased mortality risk in RA.[50] However, other factors such as infection associated with aspects such as comorbidity, including lung disease, extra-articular manifestations of disease, reduced white cell count and corticosteroid use, also contribute.[51,52]

Burden of illness

Early in disease indirect costs exceed costs due to health-care utilisation and medication (direct costs) by twofold.[53] It is also clear that informal caregivers shoulder a considerable burden in terms of forgone paid employment, leisure activity and personal health.[54] Inevitably, in a disease characterised by chronic pain, discomfort and physical impairment, the burden on individuals and families is increased. Medication costs, especially in those treated with biologic agents such as TNF inhibitors, account for a majority of the direct costs of RA.[55] Some drug intervention studies have shown reduced work absence with aggressive treatment strategies,[56] although only one-third of employed patients cease work because of disease and, unsurprisingly, manual workers are much more likely to stop work.[57] It is estimated that the total costs of RA to the UK economy is between £3.8 and £4.8 billion.[6]

Current service provision

Services for patients with RA have been reviewed in detail in a recent report by the NAO.[6] Diagnosis and management of RA is led primarily by consultant rheumatologists employed by acute hospital trusts. People who may have RA often seek help late and may suffer owing to delayed treatment and referral. There are around 460 consultant rheumatologists in England, giving a ratio of 1 : 100,000 rheumatologists per head of population (the ratio in Wales is 1 : 106,000). Consultants are supported by specialist nurses and the NAO census identified 377 specialist rheumatology nurses in England. Considerable variations and deficiencies in service provision were identified by the NAO. Specific recommendations for improving services were made by the NAO in the following areas:

- timely diagnosis and treatment
- better integration between primary and secondary care services
- improved holistic care including strategies to improve self-management and providing support for maintaining employment.

Description of the technologies

Five intervention technologies are considered in this report. Three are TNF inhibitors (ADA, ETN and IFX), and one each a T-cell costimulation modulator (ABT) and a selective CD20 B-cell depleting agent (RTX). The technologies are described below. Licensed indications and relevant NICE guidance are detailed in *Table 1*.

Tumour necrosis factor inhibitors

Adalimumab

Adalimumab is a recombinant monoclonal antibody, made from human peptide sequences, which neutralises the biological functions of tumour necrosis factor alpha (TNFα) by binding to TNF cell-surface receptors. ADA is licensed for use in RA, juvenile idiopathic arthritis, psoriasis, psoriatic arthritis, ankylosing spondylitis and Crohn's disease.

TABLE 1 European Union licensed indications related to RA for the five technologies and relevant NICE guidance

Drug	Indications and population	Doses and routes of administration	Synopsis of relevant NICE guidance
ABT	Moderate-to-severe RA – in combination with MTX. Patients with insufficient response to DMARDs including at least one TNF inhibitor	Intravenous infusion over 30 minutes. Dose according to weight, range 500–1,000 mg. Infusions at 0, 2 and 4 weeks followed by 4-weekly maintenance infusions indefinitely	TA141 Not recommended
ADA	Moderate-to-severe RA – in combination with MTX (unless MTX inappropriate). Patients with insufficient response to DMARDs including MTX	Subcutaneous injection of 40 mg every other week indefinitely. Dose may be increased to 40 mg weekly if patients experience a decrease in their response (monotherapy)	TA130 and TA36 (for ADA, ETN and IFX) DAS28 score of > 5.1 measured on at least two occasions, 1 month apart Previous trial of two DMARDs including MTX (unless contraindicated) necessary Normally used in combination with MTX – unless intolerant or inappropriate when monotherapy with ADA and ETN may be given Only continue after 6 months if DAS28 improves by > 1.2 Alternative TNF inhibitor may be considered if treatment is withdrawn due to an adverse effect before the initial 6-month assessment of efficacy Dose escalation above licensed starting dose is not recommended TA36 does not recommend the consecutive use of TNF inhibitors. This recommendation is not reproduced in the NICE RA guideline. TA130 does not report on consecutive use
ETN	Moderate-to-severe RA – monotherapy or in combination with MTX in those with an inadequate response to DMARDs. Patients with severe RA not previously treated with MTX may also be treated	Subcutaneous injection of 25 mg twice a week or 50 mg weekly given indefinitely	
IFX	Moderate-to-severe RA – in combination with MTX (unless contraindicated) in those with an inadequate response to DMARDs. Patients with severe RA not previously treated with MTX or other DMARDs may also be treated	Intravenous infusion over 2 hours at a dose of 3 mg/kg at 0, 2 and 6 weeks followed by 8-weeky maintenance infusions indefinitely. If response lost or inadequate, stepwise increases in dose by 1.5 mg/kg every 8 weeks may given up to a maximum of 7.5 mg/kg. Alternatively, dosing at 3 mg/kg may be given as frequently as 4-weekly	
RTX	Severe RA in combination with MTX in patients who have had an inadequate response or intolerance to other DMARDs including one or more TNF inhibitor	Intravenous infusion given as a course of two infusions (1,000 mg each) 2 weeks apart. Further infusions may be given but a precise limit is not given. Repeat course of treatment must not be given within 16 weeks	TA126 Use in combination with MTX in severe RA not responding to DMARDs including at least one TNF inhibitor Continue only if DAS28 improves by > 1.2 Repeat courses to be given no more frequently than every 6 months

TA, technology appraisal.

Etanercept

Etanercept is a combination protein consisting of the extracellular portion of two TNFα receptors (75-kDa TNF receptors) combined with a human fragment crystallisable (Fc) portion of the human immunoglobulin G1 (IgG1). ETN inhibits TNFα activity by binding soluble and cell-bound TNFα with high affinity and by competing with natural TNFα receptors. ETN is licensed for use in RA, psoriatic arthritis, psoriasis and ankylosing spondylitis.

Infliximab

Infliximab is a recombinant chimeric human–murine monoclonal antibody that binds soluble and membrane-bound TNFα thereby, inhibiting the functions of TNFα. IFX is licensed for use in RA, Crohn's disease, ulcerative colitis, ankylosing spondylitis, psoriatic arthritis and psoriasis.

Other tumour necrosis factor inhibitors

Certolizumab pegol has been granted a marketing authorisation in the European Union (EU) for the treatment of moderate-to-severe RA. It is administered by subcutaneous injection. Certolizumab pegol was the subject of a separate NICE single technology appraisal (STA),[58] with guidance published in February 2010. Golimumab is currently being assessed by the European Medicines Agency. A positive opinion has been given for the granting of marketing authorisation in RA. Golimumab has been referred to NICE, but the appraisal has been suspended because the manufacturer is not in a position to submit evidence to NICE.

Special precautions for use of tumour necrosis factor inhibitors

TNFα is a key component of host defence against *Mycobacterium tuberculosis* (MTB), especially by forming granulomas and preventing dissemination of mycobacteria.[59,60] Inhibition of TNFα increases the risk of MTB and other granulomatous diseases, such as those due to *Listeria monocytogenes* (a bacterium associated with food-borne diseases) and *Histoplasma capsulatum* (a fungus which, in endemic areas, causes lung disease in people with a compromised immune system). Recommendations for screening patients for tuberculosis (TB) before treatment have been published.[61] In the UK this is done most commonly by taking a medical history focusing on TB and a pre-treatment chest radiograph. Some centres also perform a tuberculin skin test,[62] although interpretation of such tests is complicated by the UK's previous vaccination programme for TB prevention and also the fact that many patients with RA respond poorly to tuberculin (possibly because of current immunosuppressive therapy but also because of the disease).[63]

Routine monitoring of blood tests is not necessary for patients taking TNF inhibitors, but is needed for concomitantly used DMARDs such as MTX. TNF inhibitors can induce anti-nuclear and anti-double-stranded DNA antibodies in the blood of some patients treated with TNF inhibitors. These antibodies are associated with systemic lupus erythematosus (SLE), a potentially serious rheumatic disease. Cases of drug-induced SLE have been reported with TNF inhibitors, but are rare.[64]

Other technologies

Rituximab

Rituximab is a chimeric monoclonal antibody which binds the CD20 cell surface marker found on B lymphocytes and depletes these cells. CD20 occurs on normal and malignant B lymphocytes (as in non-Hodgkin's lymphomas). Normal plasma cells, an important component of host defence, and haematopoietic stem cells do not carry CD20. RTX is licensed for use in RA, non-Hodgkin's lymphoma and chronic lymphocytic leukaemia. A small number of cases of progressive multifocal leucoencephalopathy, a rare but usually fatal demyelinating brain disease, have been reported in RA patients following RTX treatment.[65]

Abatacept

Abatacept is a fusion protein consisting of CTLA-4 (cytotoxic T-lymphocyte-associated antigen-4) linked to a modified Fc portion of the human IgG1. ABT works by blocking activation of certain populations of T lymphocytes. ABT is currently licensed for use only in RA.

Tocilizumab

Tocilizumab was the subject of a separate NICE STA,[66] with guidance published in August 2010. This guidance is likely to have a key impact on the treatment pathways considered in this review. TOC is a humanised monoclonal antibody that inhibits the activity of the cytokine interleukin-6 (IL-6). In the EU it is licensed for use only in moderate-to-severe RA patients who are intolerant, or have responded inadequately, to one or more DMARDs or TNF inhibitors. The drug is recommended for use in combination with MTX, but may be used alone in patients intolerant of MTX or for whom it is contraindicated. TOC is given by intravenous (i.v.) infusion over 1 hour once a month indefinitely.

Disease-modifying antirheumatic drugs, biologics, treatment sequences and combinations

Rheumatoid arthritis is characterised, in many patients, by an excellent initial response to a DMARD with subsequent loss of response with time. Most randomised trials are of a relatively short duration (typically less than 12 months) and do not study a treatment pathway. Trials of DMARDs sequences are increasingly common.[25,67,68] Remission is possible in early disease with MTX alone or in combination with other agents such as sulfasalazine, HCQ, CyA and TNF inhibitors. The optimal sequence is yet to be determined, and perhaps the choice of drug is not relevant, but the key to successful management appears to be regular patient review with a focus on optimal disease control.

The NICE RA guidance is consistent with this approach, although recent trials indicate that early use of MTX in combination with a TNF inhibitor provides better outcomes.[25,69] NICE recommends that TNF inhibitors are used only in those not responding to MTX and another DMARD. Delayed addition of a TNF inhibitor need not necessarily compromise medium-term outcomes[23,25,69] and may be justified on health-economic grounds.

What steps should be taken when a first TNF inhibitor and several DMARDs including MTX fail? This technology assessment report sets out to examine clinical effectiveness and cost-effectiveness evidence from available randomised controlled trials (RCTs), observational studies and economic evaluations. A small survey conducted as part of this technology assessment on a convenience sample of consultant rheumatologists in the West Midlands indicated considerable variability in approach for patients who fail a first TNF inhibitor. The most common suggested approaches were to consider a second TNF inhibitor or RTX (in combination with MTX). Further details of this survey can be found in *Appendix 11*.

There are many and increasing permutations of treatment sequences. Combinations of biologic agents are not licensed and where combinations have been tried there is an increased risk of serious infections. Potential drug toxicity of newly licensed agents is an important unknown. Other considerations include practical matters to do with drug delivery such as i.v. or subcutaneous administration and availability of infusion facilities. Patients with RA tend to be risk averse[70] and strategies mandating targeted disease control in late 'stable' RA are commonly resisted by doctors and patients.[71] However, in those with active and progressive disease new therapies are needed. This review seeks to explore some aspects of these uncertainties as determined by a protocol agreed with NICE and interested parties.

Degree of diffusion and anticipated costs

The number of RA patients currently being treated with TNF inhibitors is unknown. By July 2009, 12,626 patients who started treatment with a TNFα inhibitor were registered with the British Society for Rheumatology Biologics Registry (BSRBR). This register has stopped recruiting patients with RA starting ADA, ETN and IFX. So far 2,876 (23%) have ceased taking the first prescribed TNFα inhibitor and switched to a second TNFα inhibitor [1,881 switched owing to the lack of efficacy and 995 because of an an adverse event (AE)]. Of these the mean and maximum observed duration of treatment with a second TNFα are currently 18 months and 64 months, respectively. By August 2009 the BSRBR had registered 442 patients treated with RTX from a target of 1,100.[72]

The drug costs of biologic agents are similar for the agents given by subcutaneous injection at around £9,000 per annum. Costs of i.v. administered drugs vary depending on patient weight and frequency of treatments courses (with RTX). Likely drug costs for these agents range between £7,000 and £10,000 per annum.

Chapter 2

Definition of the decision problem

Decision problems

According to the final scope issued by NICE for this technology appraisal, the decisions to be made are:

- *Decision problem 1*: whether there are significant differences in clinical effectiveness and cost-effectiveness between ADA, ETN, IFX, RTX and ABT (referred to as 'the interventions' hereafter), when used within their licensed indications in adults with active RA who have had an inadequate response to a first TNF inhibitor prescribed according to current NICE guidance.
- *Decision problem 2*: whether the interventions are clinically effective and cost-effective compared with previously untried conventional DMARDs (such as LEF and CyA).
- *Decision problem 3*: whether the interventions are clinically effective and cost-effective compared with other biologic agents (including TOC, golimumab and certolizumab pegol).
- *Decision problem 4*: whether the interventions are clinically effective and cost-effective compared with supportive care (including conventional DMARDs to which patients have had inadequate response).
- *Decision problem 5*: whether the clinical effectiveness and cost-effectiveness of the interventions differ significantly between certain subgroups of patients (see *Definition of the interventions*).

The assessment report set out to address these decision problems as they apply to potential patient pathways in the UK. The nature of evidence and the timelines for this technology appraisal constrain the focus of the assessment report to key clinically relevant questions.

Definition of the interventions

The interventions being considered are:

- *Adalimumab*: a TNF inhibitor administered by subcutaneous injection and usually prescribed in combination with MTX, except in cases where MTX is not tolerated or is contraindicated.
- *Etanercept*: a TNF inhibitor administered by subcutaneous injection in combination with MTX, except in cases where MTX is not tolerated or is contraindicated.
- *Infliximab*: a TNF inhibitor administered by i.v. infusion in combination with MTX.
- Rituximab: a monoclonal antibody directed at CD20+ B cells, administered by i.v. infusion in combination with MTX.
- *Abatacept*: a T-cell costimulation modulator, administered by i.v. infusion in combination with MTX.

Population and relevant subgroups

The population being considered is adults with active RA who have had an inadequate response to a first TNF inhibitor.

Potentially relevant subgroups are numerous and include:

- patients having had primary or secondary (had initial response, but subsequently lost the response over time) failure of response to the first TNF inhibitor or having withdrawn from the first TNF inhibitor mainly owing to adverse effects
- subgroups defined by autoantibody status [e.g. presence or absence of RF and/or anti-cyclic citrullinated peptide (anti-CCP) antibodies]
- subgroups defined by different doses of the intervention (within licence)
- patients with comorbidities for which some treatments may be contraindicated (e.g. heart failure).

The specific subgroups examined in the effectiveness review of this report were determined in light of available evidence and in consultation with clinical experts. Subgroups were not considered in the economic modelling as compelling evidence of differential effectiveness between subgroups was lacking from the effectiveness review.

Clarification of population of interest

The NICE guidance states that an alternative (second) TNF inhibitor may be considered for patients in whom treatment is withdrawn because of an AE before the initial 6-month assessment of efficacy. This group of patients (withdrawal because of an early AE) is strictly speaking outside the remit of this technology appraisal and should ideally be excluded from the technology assessment. However, in practice, the reason for the withdrawal of a TNF inhibitor may not be clear-cut as a decision to withdraw may be related to both efficacy and adverse effects (and the balance of risk and benefit for the patient).

Relevant comparators

Potential comparators include:

- supportive care (including corticosteroids and ongoing or reinstated conventional DMARDs, such as MTX, sulfasalazine to which the patients have had inadequate response previously)
- conventional DMARDs which have not been tried prior to trying a TNF inhibitor for example AZA, CyA and GST, either as monotherapy or combined with other DMARDs or corticosteroids
- biologic agents including TOC, golimumab and certolizumab pegol
- the interventions being considered compared with each other.

Clarification of comparators

The assessment report focuses on key clinically relevant questions, including, where data allow, comparing each of the interventions with supportive care and comparing each of the interventions against each other. This was based on the following considerations:

- The majority of patients considered in this technology appraisal may have already had inadequate response to at least two conventional DMARDs, including MTX tried for an adequate length of time and at adequate doses, as indicated in the current NICE guidance.

These DMARDs may still be continued in the comparator (and intervention) arm(s) of trials in patients who have responded inadequately to these options. In such cases continued use of these DMARDs was regarded as supportive care rather than as a credible alternative treatment option. Therefore, a clear distinction was made between conventional DMARDs depending on whether the patients had tried them before and if there was a history of inadequate response to the DMARD tried.

- Only conventional DMARDs to which the patients have not had inadequate response or have not tried were to be regarded as separate comparators. The evidence for use of conventional DMARDs in patients who have failed to respond to TNF inhibitors was expected to be very limited.
- Although conventional DMARDs which are continued and to which the patients had an inadequate response were regarded as supportive care, subgroup analysis was considered (where relevant and evidence permits) to assess whether the presence or absence of these (failed) DMARDs in the control and intervention groups influenced the estimated treatment effects of the interventions.
- Tocilizumab, golimumab and certolizumab pegol were potentially relevant comparators. These drugs are not yet available in the UK, but all are (or are potentially) the subject of STAs by NICE. The inclusion of these three drugs in the final scope as comparators means that there were no formal submissions from their manufacturers for this technology appraisal. This may have had implications with regard to the acquisition of evidence for these comparators. It was proposed that TOC, golimumab and/or certolizumab pegol could have been reviewed in the assessment report as a comparator if marketing authorisation of the technology was obtained before the submission of the protocol for this assessment report. This condition was not met.

Relevant outcomes

Key outcomes considered appropriate to the decision problem were:

- withdrawals (with reason)
- treatment response (ACR)
- disease activity (DAS)
- physical function (HAQ)
- joint damage/radiological progression
- pain
- fatigue
- serious AEs (including death)
- other AEs potentially associated with treatment
- health-related QoL (HRQoL).

Key issues

Key issues have been mentioned, where relevant, earlier in this section and also in the background section of this report.

Further key issues predominantly concern the limited availability of evidence from controlled trials and the impact this has on the assessment of clinical effectiveness and cost-effectiveness of each of the interventions compared with the potential comparators (and the other interventions), and the ability to identify relevant subgroups in whom the technologies are more or less beneficial.

Place of the intervention in the treatment pathway(s)

Based on the final scope, the interventions are to be used when patients have had an inadequate response to a TNF inhibitor.

Overall aims and objectives of assessment

The overall aims and objectives were to address the decision questions outlined in section *Decision problems*. These aims were to be achieved by:

- A systematic review of RCTs of the efficacy, tolerability and safety of ADA, ETN, IFX, RTX and ABT for the treatment of RA in adults who have had an inadequate response to a first TNF inhibitor.
- As the volume of RCT evidence was expected to be relatively small, relevant non-randomised comparative studies and uncontrolled studies were also reviewed.
- A systematic review of published studies on the cost and cost-effectiveness of the technologies in the treatment of RA in adults who have had an inadequate response to a first TNF inhibitor.
- A review of economic evaluations included in any manufacturer's submissions (MSs) for this appraisal.
- A focused, model-based economic evaluation of the cost-effectiveness of the technologies from the perspective of the UK NHS.

Chapter 3

Assessment of clinical effectiveness

Methods for reviewing effectiveness

Search strategy

The following resources were searched for relevant studies:

- Bibliographic databases: Cochrane Library (CENTRAL) 2009 Issue3, MEDLINE (Ovid) 1,950 to July week 1 2009, MEDLINE In-Process & Other Non-Indexed Citations (Ovid) 13 July 2009, EMBASE (Ovid) 1980–2009 week 28. Searches were based on index and text words that encompassed the condition, RA, and the interventions ADA, IFX, ETN, RTX and ABT.
- Citations of included studies were examined.
- Reference lists of identified systematic reviews were checked.
- Further information was sought from contacts with experts.
- Research registries of ongoing trials including the National Institute for Health Research (NIHR) Clinical Research Network Portfolio Database, Current Controlled Trials and Clinical Trials.gov using terms for the particular drugs.
- Manufacturer submissions.

The searches were not limited by date of publication or language.

Search strategies can be found in *Appendix 2*.

Study selection

All articles identified in the searches were imported into a REFERENCE MANAGER database (REFERENCE MANAGER v.11, Thomson ResearchSoft). Duplicate entries were allowed to be removed by the inbuilt feature in REFERENCE MANAGER and removed when encountered by reviewers. Titles and abstracts were independently checked for relevance based on the population and intervention by two reviewers. If articles were considered relevant by at least one of the reviewers a full paper copy was ordered.

Full papers were assessed for relevance by two independent reviewers using an inclusion/exclusion checklist (see *Appendix 6*) based on the following criteria:

- population: a majority of adults with active RA who have had an inadequate response to a TNF inhibitor
- intervention: ADA, ETN, IFX, RTX, or ABT
- outcomes: clinical outcomes related to efficacy, safety or tolerability
- study design: primary study (except case reports) or a systematic review
- study duration: at least 12 weeks
- participant numbers: for non-randomised studies – at least 20 patients in one arm.

Disagreements were resolved by discussion with the involvement of a third reviewer when necessary.

Conference abstracts were not sought. If they were identified as relevant in the first stage of study selection, an attempt was made to match them with journal publications. If this was not possible, contact with authors was not attempted owing to time constraints and they were not included in the analysis.

A list of excluded studies and the reason for exclusion were recorded (see *Appendix 4*).

Included systematic reviews were not themselves systematically reviewed, but were utilised to identify further primary studies.

Additional references identified from systematic reviews or industry submissions were entered into the REFERENCE MANAGER database. The same process was applied to additional the references as to the references identified from initial searches.

Data extraction

Data were extracted into a standard form (see *Appendix 8*) for all included studies by one reviewer. A second reviewer checked the accuracy of the extracted information. Disagreements were resolved by consensus or by referral to a third reviewer if necessary.

Information regarding study design and characteristics of study participants was extracted. Data on the following outcomes were sought from included studies:

- treatment withdrawal (and reasons for withdrawal)
- ACR20, ACR50, ACR70
- disease activity (e.g. DAS28 or DAS)
- physical function (e.g. HAQ)
- joint damage/radiological progression (measured by a scoring system)
- pain
- fatigue
- extra-articular manifestations of the disease
- serious AEs (including death)
- other adverse effects potentially associated with treatments
- HRQoL.

Data for any outcomes other than those listed above were also extracted if they were considered relevant to this report.

Additional data from industry submissions were extracted by only one reviewer owing to time constraints.

Quality assessment

The quality of included studies was assessed independently by two reviewers. Any disagreements were resolved by discussion and if necessary a third reviewer was consulted.

For randomised trials the following criteria were considered:

- Randomisation: whether allocation was truly random. Randomisation using a computer or a random number table was considered adequate, whereas the use of alternation, case record numbers, or dates of birth and day of the week was considered inadequate.
- Allocation concealment: whether allocation concealment was adequate. Any of the following methods was considered adequate: centralised (e.g. allocation by a central office unaware of subject characteristics) or pharmacy-controlled randomisation; pre-numbered or coded identical containers which are administered serially to participants; on-site computer system

combined with allocations kept in a locked unreadable computer file that can be accessed only after the characteristics of an enrolled participant have been entered; or sequentially numbered, sealed, opaque envelopes.
- Blinding: use of blinding and who was blinded (patients, study investigators/outcome assessors, data analysts).
- Patients withdrawn: what was the percentage of patients withdrawn from the study?
- Intention-to-treat (ITT) analysis: whether ITT analysis was used.

For non-randomised studies the following criteria were considered:

- Study design: if the study was controlled or uncontrolled, prospective or retrospective.
- Inclusion criteria: if inclusion criteria were clearly stated.
- Consecutive patients: if consecutive patients were included in the study.
- Patients withdrawn: what was the percentage of patients withdrawn from the study?

The results of quality assessments are reported in relevant sections of the report.

Data analysis/synthesis

Outcomes of interest

Selected outcomes of interest were specified in the review protocol, based upon the final scope issued by NICE for this technology appraisal. These were:

- treatment withdrawal (and reasons for withdrawal)
- ACR20, ACR50, ACR70
- disease activity (e.g. DAS28 or DAS)
- physical function (e.g. HAQ)
- joint damage/radiological progression (measured by a valid scoring system)
- pain
- fatigue
- extra-articular manifestations of the disease
- serious AEs (including death)
- other adverse effects potentially associated with treatment
- HRQoL.

Handling of data and presentation of results

Comparisons with supportive care

Studies were considered to compare interventions with supportive care if they:

- had an arm receiving supportive care
- had a placebo arm.

Owing to the paucity of evidence from controlled studies of TNF inhibitors, evidence from uncontrolled studies (i.e. single-group before-and-after studies) is also considered in this section.

Studies were considered separately for each of the interventions. In addition, TNF inhibitors were discussed together as a class of drugs. Results were presented in figures and discussed in the main text of the report for the following outcomes:

- withdrawals (for any reason, owing to the lack of efficacy and owing to AEs)
- ACR20, ACR50 and ACR70
- DAS
- European League Against Rheumatism (EULAR) response
- HAQ

- QoL
- joint damage
- serious AEs
- infections and serious infections
- injection/infusion reaction.

For other outcomes only figures were created, and these can be found in *Appendix 10*.

Dichotomous measures data are presented as relative risks (RRs) (for RCTs) and percentages (for other study designs). For continuous outcomes, mean differences (for RCTs) and means (for other study designs) were used.

Where available, data were analysed for 3, 6, 9, 12, etc. months' duration of follow-up. They were assumed to be 3-month data if they were collected between 3 and 4 months from the initiation of treatment, 6-month data if they were collected between 5 and 7 months from the initiation of treatment. If more than one estimate was available for a time interval, the value nearest to the assumed follow-up was used.

Pooling of results was not attempted for the assessment of effectiveness of individual technologies because the majority of included studies had no control group and there was substantial methodological and clinical heterogeneity between included studies. Given the relatively small number of patients that can be analysed in subgroup analyses, some pooling of data using a random-effects model was attempted. The results were presented with I^2 statistics mainly for demonstrating consistency of findings between studies (see *Subgroup analyses*).

Comparisons with newly initiated and previously untried conventional disease-modifying antirheumatic drugs

No studies were identified and therefore analyses were not undertaken.

Comparisons with other biologic agents

No studies were identified and therefore analyses were not undertaken.

Comparisons between technologies (head-to-head comparisons)

No studies were identified and therefore direct comparisons were not undertaken.

Indirect comparison (IC) was undertaken when data were available from RCTs. It was conducted using the method by Bucher *et al.*[73] The results of the analyses were presented in tabular format.

Subgroup analyses

The following subgroups were specified in the review protocol:

- patients having withdrawn from the first TNF inhibitor owing to the lack of response (primary failure), loss of response (secondary failure) or AEs/intolerance
- subgroups defined by autoantibody status (e.g. presence or absence of RF or anti-CCP antibodies)
- subgroups defined by different doses of the intervention (within licence)
- patients with comorbidities for which some treatments may be contraindicated (e.g. heart failure).

No subgroup data concerning the last two categories (varied doses; comorbidities) were identified, and thus no subgroup analysis was performed for these. Subgroup analyses relating to the reasons of withdrawal from the first TNF inhibitor were carried out as two separate comparisons:

1. withdrawal owing to lack of response versus withdrawal due to loss of response
2. withdrawal owing to lack of efficacy (which includes both lack of response and loss of response) versus withdrawal due to AEs/intolerance.

In addition to the above, subgroup data in relation to the identity of the first TNF inhibitor which the patients received before discontinuation and the number of prior TNF inhibitor(s) that the patients had tried before switching were reported in some studies. These were considered potentially of clinical relevance and thus subgroup analyses on these were also performed where data were available [commercial-in-confidence information (or data) removed].

Ongoing studies

Ongoing primary studies were identified in the searches. They were not included in the systematic review, but discussed in *Ongoing studies*.

Assessment of publication bias

All manufacturers of the interventions provided a list of all company-sponsored RCTs and other non-randomised or uncontrolled studies that are relevant for this appraisal. Requests of clarification of trial data that are potentially available but not reported in published papers were also made to the manufacturers of RTX and ABT.

The number of relevant studies for individual technology was too small to allow a formal assessment of publication bias.

Sensitivity analyses

The protocol specified that if evidence permits sensitivity analyses may be carried out taking into account the following factors:

- quality measures of studies such as blinding and randomisation
- factors associated with the characteristics of the study population
- factors associated with study design such as study duration and drug doses
- exclusion of data supplied as commercial/academic in confidence.

However, sensitivity analyses were not performed as no pooling of study results was undertaken.

Changes to the original protocol

During the study selection process, several potentially relevant studies including mixed proportion of patients with or without prior treatment with a TNF inhibitor were identified. No criterion relating to inclusion or exclusion of these studies was specified in the original protocol. It was agreed by consensus within the project team that studies that included less than 50% of patients with RA who have failed a TNF inhibitor were excluded, unless results from these patients were described separately and the number of these patients was greater than or equal to 20.

This report contains reference to confidential information provided as part of the NICE appraisal process. This information has been removed from the report and the results, discussions and conclusions of the report do not include the confidential information. These sections are clearly marked in the report.

Results: quantity and quality of research available

The searches resulted in the identification of 10,281 records and an additional 17 were identified from industry submissions and 15 from reference lists of included studies.

Nine relevant systematic reviews[74–82] were identified in addition to the reports conducted for previous NICE appraisals in RA. Examination of these nine reviews did not identify any further primary studies that met all the criteria for inclusion in either the clinical effectiveness or cost-effectiveness sections of this report.

Duplicates had been removed, leaving 7,486 records. Screening of the title and abstract of these articles indicated that 174 were directly relevant to the clinical effectiveness section of this report. Full paper copies of these articles were ordered. Five of them were unobtainable.[83–87] Inclusion criteria were applied to the remaining 169 articles. Of these, 113 were excluded for not meeting at least one of the inclusion criteria. Three articles were identified as conference abstracts[88–90] and, as these could not be matched to full publications, they were excluded. Details of excluded studies together with reasons for exclusion can be found in *Appendix 4*.

A flow diagram presenting the process of identification of relevant studies can be found in *Appendix 3*.

There were 35 studies described in 45 papers meeting the inclusion criteria. Five of the studies were RCTs, one was a comparative study, one was a non-randomised controlled study and 28 were uncontrolled studies [including one long-term extension (LTE) of an RCT].

A randomised study on RTX [study for understanding rituximab safety and efficacy (SUNRISE)[91]] that was not yet published in full was identified. Data from this study were requested from the manufacturer; however, the clinical study report was received too late to be included in the analyses.

Table 2 presents mapping of studies to relevant interventions and comparators.

TABLE 2 Mapping of identified studies

	Interventions (newly initiated)					
Comparators	**ADA**	**ETN**	**IFX**	**TNF inhibitors**	**RTX**	**ABT**
None[a]	Bennett 2005[92] (*n* = 26, 52 weeks) Wick 2005[93] (*n* = 27, 24 weeks) Nikas 2006[94] (*n* = 24, 52 weeks) Bombardieri 2007[95,96] (*n* = 899, 12 weeks) van der Bijl 2008[97] (*n* = 41, 16 weeks)	Haraoui 2004[98] (*n* = 25, 12 weeks) Buch 2005[99] (*n* = 207, 12 weeks) Cohen 2005[100] (*n* = 24, 13 weeks) Buch 2007[101] (*n* = 95, 12 weeks) Iannone 2007[102] (*n* = 37, 24 weeks) Laas 2008[103] (*n* = 49, >36 weeks) Bingham 2009[104] (*n* = 201, 16 weeks)	Ang 2003[105] (*n* = 24, unclear) Hansen 2004[106] (*n* = 20, unclear) Yazici 2004[107] (*n* = 21, unclear)	Gomez-Reino 2006[108] (*n* = 488, 104 weeks) Solau-Gervais 2006[109] (*n* = 70, > 13 weeks) Hjardem 2007[110] (*n* = 235, 13 weeks) Duftner 2008[111] (*n* = 109, up to 208 weeks) Karlsson 2008[112] (*n* = 337, 13 weeks) Blom 2009[113] (*n* = 197, 48 weeks)	Bokarewa 2007[114] (*n* = 48, 52 weeks) Jois 2007[115] (*n* = 20, 26 weeks)[b] Keystone 2007[116] (*n* = 158, 24 weeks) Assous 2008[117] (*n* = 50, 26 weeks) Thurlings 2008[118] (*n* = 30, 24 weeks)	ATTAIN LTE[119] (*n* = 317, < 260 weeks) ARRIVE[120] (*n* = 1,046, 24 weeks)

TABLE 2 Mapping of identified studies *(continued)*

	Interventions (newly initiated)					
Comparators	**ADA**	**ETN**	**IFX**	**TNF inhibitors**	**RTX**	**ABT**
Supportive care[c]				Hyrich 2009[121–123] (*n* = 736, > 24 weeks)	**REFLEX**[124–126] (*n* = 517, 48 weeks) **SUNRISE**[91] (*n* = 559, > 48 weeks)	**ATTAIN**[127–132] (*n* = 391, 26 weeks)
Ongoing biologics[d]			**OPPOSITE**[133] (*n* = 27, 16 weeks)			**Weinblatt 2007**[134] (*n* = 121, 52 weeks) **ASSURE**[135] (*n* = 167, 52 weeks)
Newly initiated DMARD						
ADA						
ETN						
IFX						
TNF inhibitors						
RTX				Finckh 2009[136,137] (*n* = 318, > 44 weeks)		
ABT						
TOC						
Golimumab						
Certolizumab pegol						

ARRIVE, abatacept researched in rheumatoid arthritis patients with an inadequate anti-TNF response to validate effectiveness; ASSURE, abatacept study of safety in use with other rheumatoid arthritis therapies; ATTAIN, abatacept trial in treatment of anti-TNF inadequate responders; OPPOSITE, open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept; REFLEX, randomised evaluation of long-term efficacy of rituximab in rheumatoid arthritis.

a Studies listed in this row are uncontrolled observational studies.

b Majority of patients had failed two or more TNF inhibitors.

c Including ongoing DMARDs to which the patients have had inadequate response and the control treatments in placebo-controlled trials.

d Ongoing biologics to which the patients have had inadequate response: OPPOSITE[131] – ongoing ETN; ASSURE[133] – ABT plus ongoing biologics (not specified) vs ongoing biologics (not specified).

Bold type indicates that the study was an RCT.

Weinblatt *et al.*[132] and ASSURE[133]: with ongoing biologic therapy in both arms; SUNRISE has not yet been published.

The assessment of effectiveness of the technologies is reported below in six sections, one for each of the technologies and one for TNF inhibitors as a class (see *Effectiveness of the technologies compared with supportive care*). Studies directly comparing the technologies and ICs are reported in *Evidence from comparative studies* and *Indirect comparisons* sections respectively.

Effectiveness of the technologies compared with supportive care

This section describes evidence relating to each of the technologies compared with supportive care, which includes treatments received by the placebo group in placebo-controlled trials and ongoing conventional DMARDs or biologics to which the patients had had inadequate response. Owing to the paucity of evidence from controlled studies for TNF inhibitors, evidence from uncontrolled studies (i.e. single-group before-and-after studies) is also considered in this section.

Adalimumab

Overview of evidence

Five studies in six publications[92–97] met the inclusion criteria. No RCT was found. Four studies had comparator arms in which the patients were TNF inhibitor naive.[92–94,96] These arms were excluded here. One of the four studies[93] also had a small comparator arm of nine patients, which did not meet the inclusion criteria of this report of greater than or equal to 20 patients for an arm to be included; thus, data from this arm were excluded.

One multicentre study was conducted in 12 countries, 11 of which were European, including the UK. Other studies were conducted in the UK, Sweden and Greece. It was unclear in which country one of the studies was conducted.

Sample sizes were small, ranging from 24 to 41 patients, that are relevant to the review in four studies; in one study there were 899 patients. Patients included all had previous treatment with either one or two TNF inhibitors, most frequently IFX. Reasons for switching TNF inhibitors were lack of efficacy only in one study,[93] lack of efficacy or intolerance in two studies[96,97] and lack of efficacy or AEs in two studies.[92,94] Details on ADA treatment were not reported in one study; in all the other studies ADA was given 40 mg subcutaneously every other week. Study duration ranged from 12 weeks to over 1 year. Further details are outlined in *Table 3*.

Patient characteristics

Data on patient characteristics can be found in *Table 4*. Characteristics of the patients included in the five studies varied in some aspects:

TABLE 3 Adalimumab: characteristics of included studies

Study	Country	Design	Reason for switching (*n*)	Prior TNF inhibitors (*n*)	Treatment arms (no. of patients)	Duration of follow-up	Comments
Bennett 2005[92]	UK	Uncontrolled prospective	Primary (8) and secondary (13) failure, AEs, other	IFX, ETN, anakinra (1)	ADA, (26)	> 52 weeks	Primary and secondary failures – all IFX
Wick 2005[93]	Sweden	Uncontrolled retrospective	Secondary failure	IFX (1)	ADA, (27)	3, 6 months	
Nikas 2006[94]	Greece	Uncontrolled prospective	Lack of efficacy, AEs	IFX (1)	ADA, (24)	12 months	Possibly one or two active TB patients (outside study inclusion criteria)
Bombardieri 2007 (ReAct)[95,96]	Australia, Austria, Belgium, France, Germany, Greece, Italy, the Netherlands, Portugal, Spain, Switzerland, UK	Uncontrolled prospective	Primary and secondary failure, intolerance	IFX, ETN, or both (≥ 1)	ADA, (899)	12 weeks	
van der Bijl 2008[97]	Unclear	Uncontrolled prospective	Primary and secondary failure, intolerance	IFX (1)	ADA, (41)	16 weeks (follow-up to 56 weeks; treatment for and efficacy measured at 16 weeks)	Pre-existing antirheumatic therapy (in about 12 patients) was continued and remained stable until week 16

ReAct, research in active rheumatoid arthritis.

DOI: 10.3310/hta15140

TABLE 4 Adalimumab: baseline patient characteristics

Study	Number of patients/% female	Age (years), mean (SD)	RA duration (years), mean (SD)	RF positive (%)	% of patients on concomitant DMARDs and steroids	Number of previous DMARDs, mean (SD)	Number of previous TNF inhibitors, mean (SD)	HAQ, mean (SD)	DAS28, mean (SD)	TJC/SJC, mean (SD)	ESR (mm/hour), mean (SD)	CRP (mg/dl), mean (SD)
Bennett 2005[92,a]	26/87	54 (range 19–77)	NR	NR	MTX (37); LEF (3); HCQ (3); AZA (1). All above with or without low-dose prednisone	3.4 (range 2–7)	1 (IFX, ETN, anakinra)	2.07 (NR)	6.3 (NR)	NR	NR	NR
Wick 2005[93]	27/84[b]	50 (15)	NR	NR	MTX (85); steroids (NR)	2.0 (NR)	1 (all IFX)	1.39 (0.52)[c]	5.5 (1.6)[c]	Tender 8 (5)[c]; swollen 10 (5)[c]	41.7 (27.5)[c]	43.9 (45.2)[c]
Nikas 2006[94]	24/92	57 (11)	16.6 (7.0)	63	MTX (83); CyA (4); LEF (13); steroids (100)	NR	1 (all IFX)	NR	5.6 (0.8)	Given graphically only	Given graphically only	Given graphically only
Bombardieri 2007[95,96]	899/81	53 (13)	12.0 (8.0)	72	DMARDs (31), steroids (77)	5.0 (1.9)	≥1 (IFX and/or ETN)	1.85 (0.66)	6.3 (1.1)	Tender 15 (7); swollen 11 (6)	NR	NR
van der Bijl 2008[97]	41/88	55 (NR)	11.6 (7.4)	NR	One DMARD (66); steroids (NR)	NR	1 (all IFX)	1.85 (0.49)	6.1 (0.9)	Tender 6 (1); swollen 8 (5)	NR	25.1 (32.0)

NR, not reported; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count.

a Female %, mean age, previous prednisone, previous DMARDs, HAQ and DAS28 given were based on the total number of 70 patients, including those patients previously TNF inhibitor naive.

b Female % was based on the total number of 62 patients.

c SD was calculated from the standard error.

TABLE 5 Adalimumab: non-RCT quality assessment

Study	Study design	Inclusion criteria clearly defined?	Were consecutive patients included in the study?	Patients withdrew (%)	Comments
Bennett 2005[92]	Prospective uncontrolled	Yes	Yes	NR[a]	
Wick 2005[93]	Retrospective uncontrolled	No	NA	0	
Nikas 2006[94]	Prospective uncontrolled	Yes	Unclear	16.7	
Bombardieri 2007[95,96]	Multicentre, uncontrolled open-label	Yes	Unclear	9.9	
van der Bijl 2008[97]	Pilot uncontrolled open-label prospective	Yes	Unclear	26.8	

NA, not applicable; NR, not reported.

a Reported for all patients, not only those relevant to the report.

TABLE 6 Adalimumab: outcomes assessed in studies and reported in the main text of the report

Study	Total withdrawal	Withdrawal by reason	ACR (20/50/70)	DAS28	EULAR response	HAQ	QoL	Joint damage	Serious AEs	Infection/ serious infection	Injection/ infusion reaction
Bennett 2005[92]				✓	✓ time range	✓ time range					
Wick 2005[93]	✓	✓	✓	✓							
Nikas 2006[94]	✓	✓	✓	✓	✓					✓	✓
Bombardieri 2007[95,96]	✓	✓	✓	✓	✓	✓			✓	✓	
van der Bijl 2008[97]	✓	✓	✓	✓	✓	✓				✓	

TABLE 7 Adalimumab: outcomes assessed in studies and reported in *Appendix 10* only

Study	Other measures of disease activity	Fatigue	Pain	TJC/SJC	CRP/ESR
Bennett 2005[92]					
Wick 2005[93]					
Nikas 2006[94]			✓[a]	✓[a]	✓[a]
Bombardieri 2007[95,96]	✓				
van der Bijl 2008[97]					✓

SJC, swollen joint count; TJC, tender joint count.
a Reported graphically only.

- Where reported 81%–92% were female.
- The mean age of the patients ranged from 50 to 57 years.
- The mean RA duration ranged from 11.6 to 16.6 years, but was not reported in two studies.
- The percentage of RF-positive patients was reported only in two studies (63% and 72%).
- Concomitant DMARDs: where reported 37%–85% patients were on MTX other DMARDs included CyA (4%), leflunonide (3%–13%), HCQ (3%) and AZA (1%).
- The percentage of patients on concurrent steroids was reported in two studies and ranged from 77% to 100%.
- Where reported the mean number of previous DMARDs used ranged from 2 to 5.
- The mean number of previous TNF inhibitors was greater than or equal to 1 in the biggest study, and it was exactly 1 in all the other studies.
- The HAQ scores ranged from 1.29 to 2.07 in four studies, but were not reported in one study.
- The mean DAS28 scores were very similar, ranging from 5.5 to 6.3.
- The mean number of tender and swollen joints at baseline was reported in three studies and ranged from 6 to 15 and from 8 to 11, respectively.
- Baseline ESR was reported in only one study (41.7 mm/hour) and CRP in only two studies (25.1 mg/dl and 43.9 mg/dl).

Quality assessment

The studies were all uncontrolled; four of them were prospective and one was retrospective.[93] Criteria for patient inclusion were clearly stated in four studies; however, in three of these it was unclear whether consecutive patients were included. The highest percentage of patients withdrawn from a study was 26.8%. There were no withdrawals from the retrospective study. In general, the higher withdrawal rates occurred with the longer follow-up durations. Further details on the quality assessment of the studies are given in *Table 5*.

Results

Tables 6 and *7* show what outcomes were measured in each study. Outcomes in *Table 6* are reported and discussed in the main text of this report and those in *Table 7* are reported in *Appendix 10* only.

Withdrawals

Withdrawal rates are presented in *Figure 1*. At 3 months, the percentage of patients withdrawn was very similar in the two studies that reported this outcome (9.9% and 9.8%). No patients withdrew in a retrospective study during 6 months. Withdrawal rates reported at 1 year were 12.5% and 26.8% in the two studies that reported this outcome. The percentage of patients withdrawn owing to lack of efficacy and owing to AEs at 3 months was reported only in the biggest study and was 2.9% and 5.6%, respectively. The percentage of patients withdrawn owing to lack of efficacy and owing to AEs at 12 months was measured in two studies: 8.3% and 17.1% withdrew because of lack of efficacy and 8.3% and 14.6% withdrew because of AEs.

One study[92] reported withdrawal data based on all 70 patients, including 44 patients who received a prior TNF inhibitor as well as TNF inhibitor-naive patients; the withdrawal data were not included in this report.

ACR20 response

The ACR20 response was assessed in four studies (*Figure 2*). Two studies assessed it at 3 months and the response was achieved by around half of the patients (46% and 60%). In the other two studies, the percentage of patients who achieved ACR20 response was 70% at 6 months and 75% at 12 months.

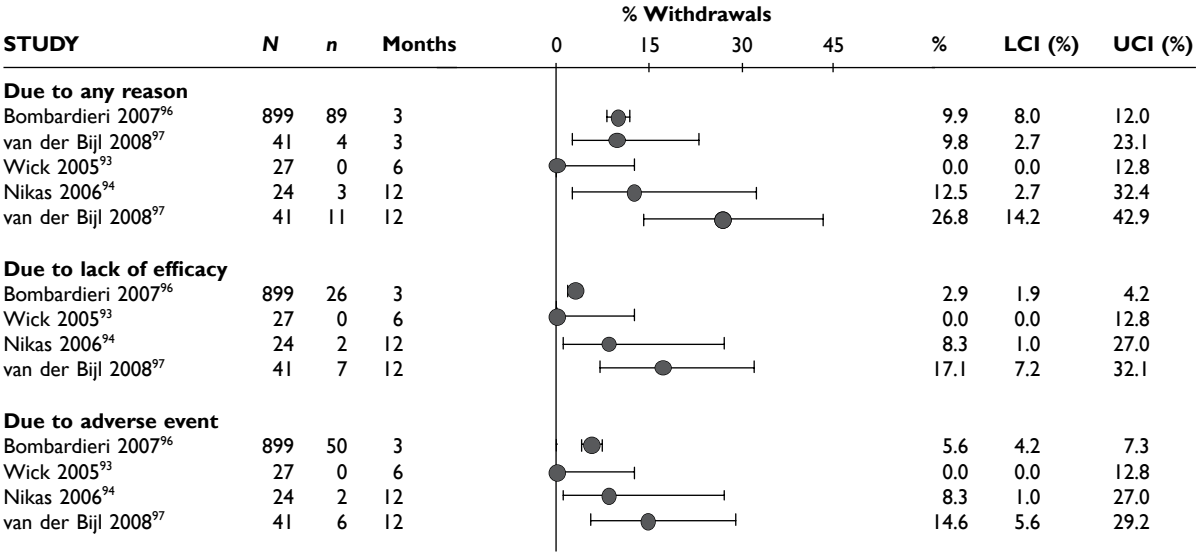


FIGURE 1 Adalimumab: withdrawals from studies by reason. LCI, lower confidence interval; UCI, upper confidence interval.

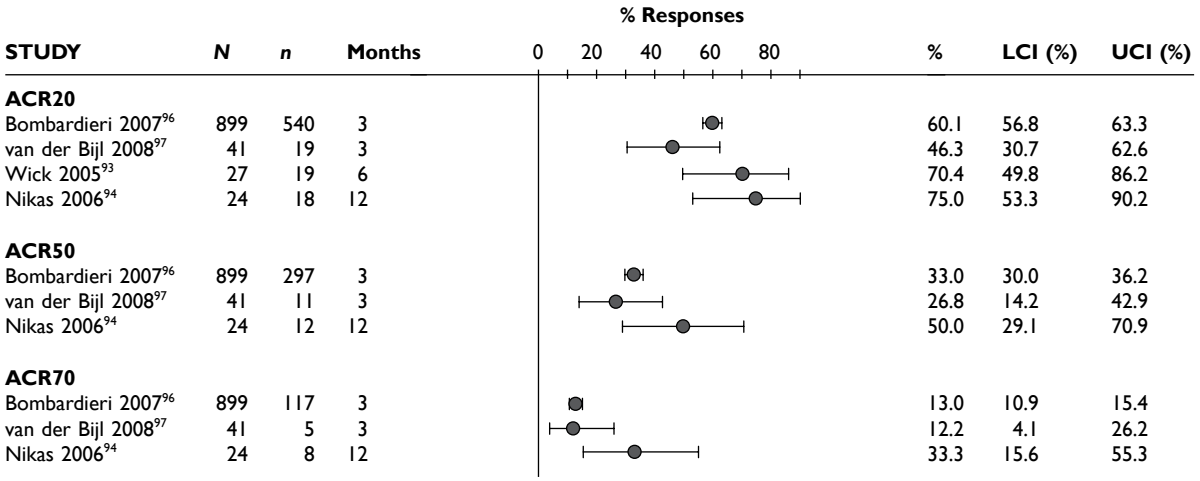


FIGURE 2 Adalimumab: ACR (20, 50, 70) responses. LCI, lower confidence interval; UCI, upper confidence interval.

ACR50 response

The ACR50 response was measured in three studies (*Figure 2*): 26.8%–33% of patients achieved ACR50 response at 3 months. When measured at 12 months in the other study, half of the patients achieved this response.

ACR70 response

The ACR70 response was measured in three studies (*Figure 2*). ACR70 response at 3 months was similar in two studies that measured this outcome (13% and 12%). ACR70 response at 12 months was reported in one study, with 33% of the patients achieving this response.

A similar pattern was seen for ACR20, ACR50 and ACR70, with a relatively higher percentage of patients achieving a response with longer duration of treatment.

DAS28

One study measured DAS28 at 3 and 6 months and another study at 12 months; the mean scores were 4.5, 4.2 and 3.2, respectively. See *Figure 3* for details. The mean changes from baseline to 3 months and to 6 months [note: in the Bennett *et al.* study[92] it was measured after mean treatment duration of 8.5 (range 1–19) months], were reported in four studies including the biggest study. They all showed that treatment with ADA significantly improved DAS28 scores (mean changes ranged from –1.30 to –1.90). See *Figure 4* for details.

EULAR response

Two studies reported EULAR response at 3 months; most of the patients had a good/moderate response (76% and 78%) and 17%–23% had a good response. The Bennett *et al.* study[92] measured EULAR response after a mean treatment duration of 8.5 months (range 1–19 months); the response rate was 65%, of whom 46% had a moderate response and 19% had a good response. See *Figure 5* for details.

Health Assessment Questionnaire

Mean change in HAQ score was reported in three studies. *Figure 6* shows that the mean HAQ score measured at 3 months in two studies, including the biggest study, and at mean 8.5 months (range 1–19 months) in the Bennett *et al.* study[92] in all cases showed a significant decrease, ranging from –0.21 to –0.48, with the largest improvement observed in the biggest study.

Joint damage

None of the studies reported this outcome measure.

Quality of life

None of the studies reported this outcome measure.

Serious adverse events

One study (the largest) reported that 18% of the patients experienced serious AEs and 13% withdrew because of AEs; none of these was lupus related or a demyelinating disorder.[96,97]

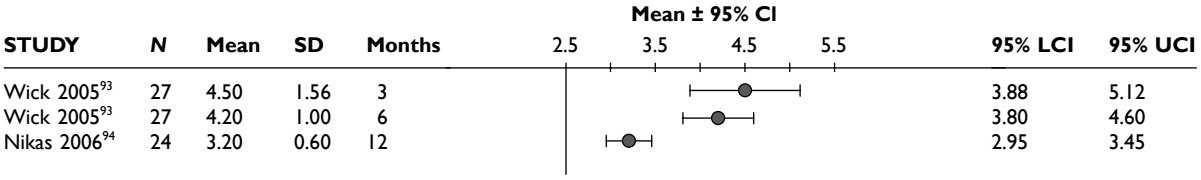


FIGURE 3 Adalimumab: DAS28 scores. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

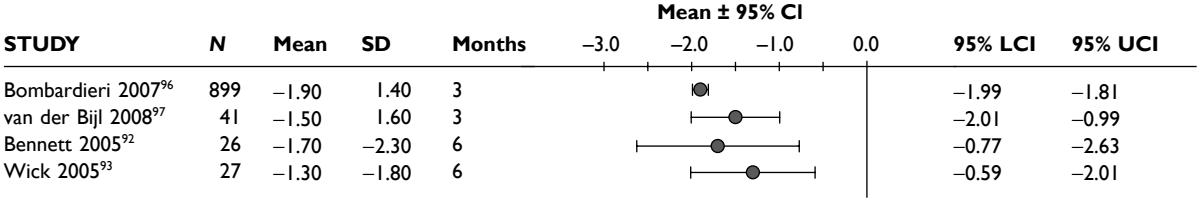


FIGURE 4 Adalimumab: mean changes in DAS28 scores. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

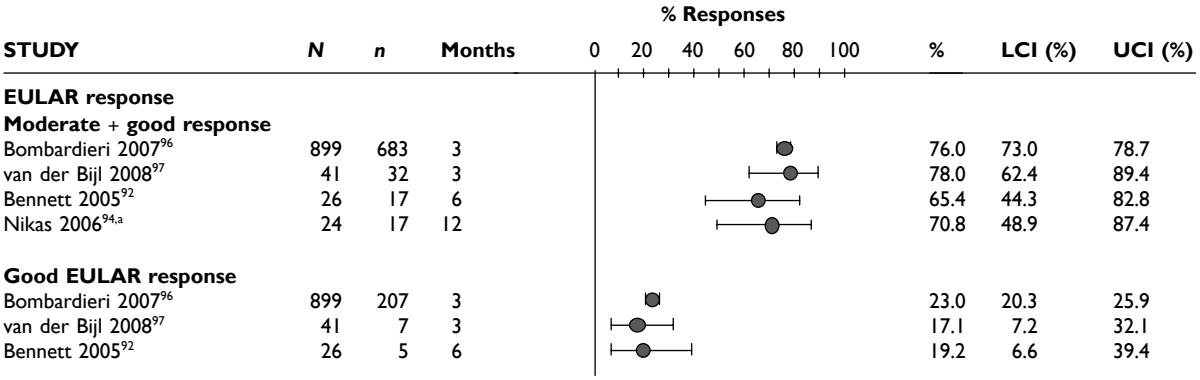


FIGURE 5 Adalimumab: EULAR response. (a) Nikas *et al.*[94] only reported 'EULAR response' without providing further detail. LCI, lower confidence interval; UCI, upper confidence interval.

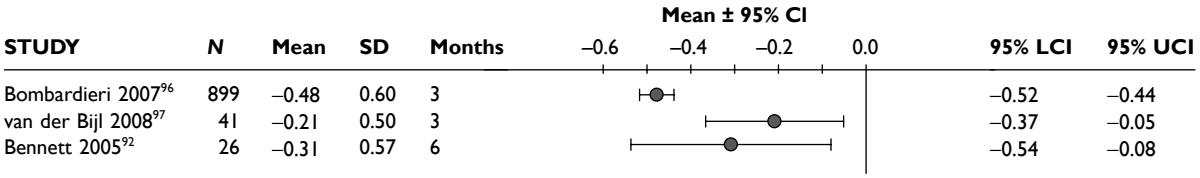


FIGURE 6 Adalimumab: mean change in HAQ scores. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

Any infections/serious infections

The largest study reported that the serious infection rate was 10.0/100 patient-years. The prevalence of TB infection was 0.4/100 patient-years in this study. In another study[97] one patient developed pulmonary TB at 11 months. In the latter study, serious infection with cellulitis was also reported in one patient. One patient in a 12-month study by Nikas *et al.*[94] had to stop the study because of herpes zoster infection; it was not reported at which time point the treatment was stopped.

Injection site reaction/infusion reaction

The largest study stated that none of the patients experienced a serious anaphylactic response during the study period of 3 months. In a 12-month study,[94] one patient had to stop the study because of an immediate hypersensitivity reaction; it was not reported at which time point it was stopped.

Summary

Five uncontrolled studies were identified for the assessment of effectiveness of ADA in comparison with standard care. Follow-up duration ranged from 3 months to over 1 year. All patients included in the studies were generally similar. The main results are summarised in *Table 8*.

Etanercept

Overview of evidence

No RCT was found. Seven uncontrolled observational studies[98–104] were identified that assessed efficacy of ETN.

In the studies by Buch *et al.*[99] and Bingham *et al.*[104] lack of efficacy was the primary reason for switching to ETN. In studies by Haraoui *et al.*,[98] Cohen *et al.*[100] and Buch *et al.*[101] patients discontinued IFX owing to a lack of efficacy or safety. In Iannone *et al.*,[102] patients had to have

TABLE 8 Adalimumab: summary of main results

Outcome	3 months	6 months	≥ 9 months
Withdrawals (%):			
■ for any reason	9.8–9.9	0	12.5–26.8
■ due to lack of efficacy	2.9	0	8.3–17.1
■ due to AEs	5.6	0	8.3–14.6
ACR20 response (%)	46.3–60.1	70.4	75.0
ACR50 response (%)	26.8–33.0	NR	50.0
ACR70 response (%)	12.2–13.0	NR	33.3
EULAR response (%):			
■ good/moderate response	76.0–78.0	65.4	70.8
■ good response	17.1–23.0	19.2	NR
remission	NR	7.7	NR
DAS28:			
■ mean change from baseline	–1.50 to –1.90 (significant improvement)	–1.30 to –1.70 (significant improvement)	NR
■ mean at time point	4.50	4.20	3.20
HAQ: mean change from baseline	–0.21 to –0.48 (significant improvement)	–0.31 (significant improvement)	NR
QoL	NR	NR	NR
Joint damage	NR	NR	NR
Serious AEs	From one study:[95,96] 18% had serious AE (no lupus-related or demyelinating disorder) and 13% withdrew because of AE	NR	NR
Any infections/serious infections	From one study:[95,96] serious infections rate 10.0/100 patient years; TB infection rate 0.4/100 patient-years	NR	From one study:[97] one patient developed pulmonary TB; one with serious cellulitis From one study:[94] one herpes zoster infection led to withdrawal
Infusion reaction	From one study:[95,96] allergic AEs 6.5/100 patient-years (no serious anaphylactic response)	NR	From one study:[94] one withdrawal because of an immediate hypersensitivity reaction

NR, not reported.

responded to prior IFX treatment but later switched to ETN due to side effects. The patient population in this study was therefore different from the other studies. In Laas *et al.*,[103] patients discontinued IFX owing to a lack of efficacy, safety or non-medical reasons. The group of patients who discontinued IFX owing to non-medical reasons (46%, 23/49) had responded well to IFX, but switched to ETN for practical reasons such as convenience (e.g. no need for hospital visit to receive infusion). Two studies[99,101] were carried out at the same centre (Leeds Teaching Hospitals) in the UK. These studies were described separately in this section although it is possible that patients included in Buch *et al.* 2005[99] were a subgroup of the cohort included in Buch *et al.* 2007.[101] The other studies were carried out in France,[100] Italy,[102] Finland[103] and the USA.[98] One study[104] was a multicentre study that enrolled patients from both the USA and Canada. The length of follow-up varied from 12 weeks to more than 9 months. Further details are provided in *Table 9*.

Patient characteristics

Full details of patients' characteristics are reported in *Table 10*. The number of patients included in the studies varied from 24 to 201. Patient characteristics differed across the seven studies:

TABLE 9 Etanercept: characteristics of included studies

Study	Country	Design	Reason for switching	Prior TNF inhibitor	Treatment arms (no. of patients)	Duration of follow-up	Comments
RCTs							
None were identified							
Non-randomised comparative studies							
None were identified							
Uncontrolled studies							
Haroui 2004[98]	USA	Uncontrolled prospective	Inefficacy and AEs	IFX	ETN (25)	12 weeks	
Buch 2005[99]	UK	Uncontrolled prospective	Inefficacy	IFX	ETN (25)	12 weeks	This study had other subgroups not relevant to this review
Cohen 2005[100]	France	Uncontrolled retrospective	Inefficacy and AEs	IFX	ETN (24)	3 months	Included an arm with 14 patients on IFX (switched from ETN)
Buch 2007[101]	UK	Uncontrolled prospective	Inefficacy and AEs	IFX	ETN (95)	12 weeks	
Iannone 2007[102]	Italy	Uncontrolled retrospective	AEs	IFX	ETN (37)	24 weeks	
Laas 2008[103]	Finland	Uncontrolled prospective	Inefficacy, AEs, non-medical reasons	IFX	ETN (49)	>9 months	Results >9 months reported but duration of follow-up unclear
Bingham 2009[104]	USA and Canada	Uncontrolled prospective	Inefficacy	IFX	ETN (201)	16 weeks	

- Where reported, the percentage of female patients ranged from 60% to 88%.
- Where reported, the mean age ranged from 49 to 57 years.
- Where reported, the mean disease duration ranged from 8.3 to 12.2 years.
- Where reported, the percentage of RF-positive patients ranged from 44% to 75%.
- Where reported concomitant DMARDs were: 88–99% MTX, other DMARDs included HCQ (9%) and sulfasalazine (5%).
- Where reported, 40%–88% of patients were receiving corticosteroids.
- Where reported, the mean/median number of previously used conventional DMARDs varied from 4.1 to 7.0.
- All the studies included patients previously treated with IFX.
- Where reported the mean baseline HAQ ranged from 0.90 to 2.16.
- The mean baseline DAS28 score ranged from 5.6 to 6.6.
- One study[102] reported baseline DAS44 (mean value was 2.7).
- Where reported, the mean number of tender and swollen joints was variable (tender: 10.0–17.8 and swollen: 8.6–14.3).
- Baseline ESR was reported only in two studies (21 mm/hour and 30 mm/hour).
- Where reported, CRP ranged from 0.6 (median) to 6.2 (mean) mg/dl.

The baseline values listed in *Table 10* for Iannone *et al.*[102] were measured 8 weeks before patients switched from IFX to ETN (while they were still responding to IFX) and thus the values may not be comparable with those from the other studies.

Quality assessment

All the seven studies were uncontrolled studies. Five were prospective[98,99,101,103,104] and two were retrospective.[100,102] Full details of the quality assessment are reported in *Table 11*. With the

TABLE 10 Etanercept: baseline patient characteristics

Study	Number of patients/% female	Age (years), mean (SD)	RA duration (years), mean (SD)	RF positive (%)	% of patients on concomitant DMARDs and steroids	Number of previous DMARDs, mean (SD)	Number of previous TNF inhibitors, mean (SD)	HAQ, mean (SD)	DAS28, mean (SD)	TJC/SJC, mean (SD)	ESR (mm/hour), mean (SD)	CRP (mg/dl), mean (SD)
Haraoui 2004[98]	25/84	50 (39)	10.0 (25.2)	NR	MTX (88), oral corticosteroid (48)	4.8 (3.7)	1	1.53 (NR)	NR	10.0 (NR)/8.6 (NR)	NR	1.7 (NR)
Buch 2005[99,a]	34/71	56 (NR)	NR	44	NR	NR	1	NR	6.4 (NR)	NR	NR	3.8–4.2[a]
Cohen 2005[100]	24/88	54 (11)	12.2 (9.6)	NR	MTX (NR)	4.1 (1.8)	1	NR	5.6 (1.1)	NR	NR	NR
Buch 2007[101]	95/NR	57 (14)	NR	71	NR	NR	1	2.16 (0.64)	6.4 (1.3)	14.0 (1.0)/9.0 (0.9)	NR	6.0 (NR)
Iannone 2007[102]	37/81	49(12)	8.3 (6.0)	75	MTX (NR), prednisone (NR)	NR	1	0.90 (NR)	2.7 (NR) (DAS44)	NR	21 (NR)	0.6[b] (NR)
Laas 2008[103]	49/88	NR	12.2 (NR)	65	MTX (NR), prednisone (88)	6.0–7.0[b]	1	NR	NR	NR	NR	NR
Bingham 2009[104]	201/60	57 (13)	9.1 (9.5)	58	MTX (99), sulfasalazine (5), HCQ (9), prednisone (40)	NR	1	1.60 (0.50)	6.6 (1.0)	17.8 (7.1)/14.3 (6.3)	30[b] (range 2–125)	6.2 (NR)

NR, not reported; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count.

a Only 25 of the 34 patients actually switched to ETN after receiving IFX. The range presented for CRP is the range of median values for the two subgroups relevant to this review (mean values not reported).

b Median value. For Laas *et al.*,[101] the range presented for the number of previous DMARDs was the range of median number of previous DMARDs among the three subgroups within the study (values for the whole study population not reported).

TABLE 11 Etanercept: non-RCT quality assessment

Study	Study design	Inclusion criteria clearly defined?	Were consecutive patients included in the study?	Patients withdrawn (%)	Comments
Haraoui 2004[98]	Uncontrolled prospective	Yes	Unclear	Unclear	
Buch 2005[99]	Uncontrolled prospective	Yes	Yes	Unclear	
Cohen 2005[100]	Uncontrolled retrospective	Yes	NA	Unclear	
Buch 2007[101]	Uncontrolled prospective	Yes	Yes	Unclear	
Iannone 2007[102]	Uncontrolled retrospective	Yes	NA	Unclear	
Laas 2008[103]	Uncontrolled prospective	No	NR	Unclear	
Bingham 2009[104]	Uncontrolled prospective	Yes	Unclear	0.5	

NA, not applicable; NR, not reported.

exception of Laas *et al.*,[103] studies stated clearly their inclusion criteria. Only Buch *et al.* 2005[99] and Buch *et al.* 2007[101] clearly stated that consecutive patients were included in the studies; this information was unclear in Bingham *et al.*[104] and Haraoui *et al.*[100] Only one study[104] reported the percentage of patients lost to follow-up (0.5%).

Results

Table 12 and *Table 13* show what outcomes were measured in each study. Outcomes in *Table 12* are reported and discussed in the main text and in *Table 13* are reported in *Appendix 10* only.

Withdrawals

Five out of seven studies reported withdrawals and the reasons for withdrawing from treatment. The percentages and reasons for withdrawing from the study after commencing ETN are shown in *Figure 7*. The percentage of patients who withdrew for any reason ranged from as low as 6.5% (at 3 months) to as high as 58.3% (at 12 months). The percentage of patients who withdrew because of AEs and lack of efficacy ranged from 0% to 16.3% and from 0% to 29.2%, respectively.

ACR20 response

The ACR20 response was assessed in four studies (*Figure 8*). The percentage of patients treated with ETN after IFX failure that achieved ACR20 response after 3 months ranged from 37.5% to 72.0%.

ACR50 response

The ACR50 response was assessed in five studies, but results from Iannone *et al.*[102] are not presented here, as explained above (*Figure 9*). The proportion of patients achieving a ACR50 response after 3 months ranged from 18.4% to 64.0%.

ACR70 response

The ACR70 response was assessed in five studies, but results from Iannone *et al.*[102] are not presented here, as explained above (*Figure 10*). The proportion of patients achieving a ACR70 response after 3 months ranged from 4.2% to 20.0%.

DOI: 10.3310/hta15140

TABLE 12 Etanercept: outcomes assessed in studies and reported in the main text of the report

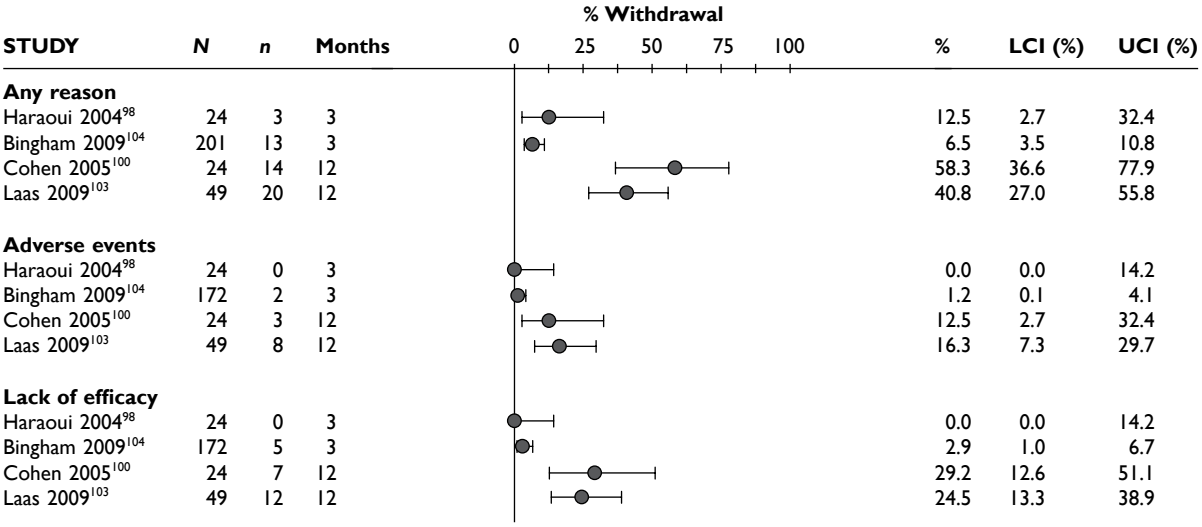
Study	Total withdrawal	Withdrawal by reason	ACR (20/50/70)	DAS28	EULAR response	HAQ	QoL	Joint damage	Serious AEs	Infection/ serious infection	Injection/ infusion reaction
Haraoui 2004[98]	✓	✓	✓			✓			✓		
Buch 2005[99]			✓								
Cohen 2005[100]	✓	✓		✓	✓					✓	
Buch 2007[101]			✓	✓	✓						
Iannone 2007[102]			✓[a]	✓		✓					
Laas 2008[103]	✓	✓		✓						✓	
Bingham 2009[104]	✓	✓	✓		✓	✓			✓	✓	

a The results for ACR50 and 70 reported by Iannone *et al.*[100] were measured against the baseline before the patients started prior IFX, not before the patients switched to ETN. The results were, therefore, not presented in this section.

TABLE 13 Etanercept: outcomes assessed in studies and reported in *Appendix 10* only

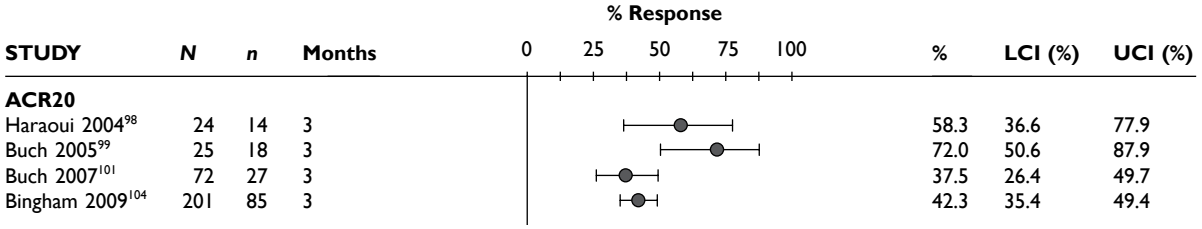
Study	Other measures of disease activity	Fatigue	Pain	TJC/SJC	CRP/ESR
Buch 2005[99]					✓
Haroui 2004[98]	✓		✓	✓	✓
Cohen 2005[100]	✓		✓		✓
Buch 2007[101]					✓
Iannone 2007[102]			✓		✓
Laas 2008[103]					✓
Bingham 2009[104]	✓		✓	✓	✓

SJC, swollen joint count; TJC, tender joint count.



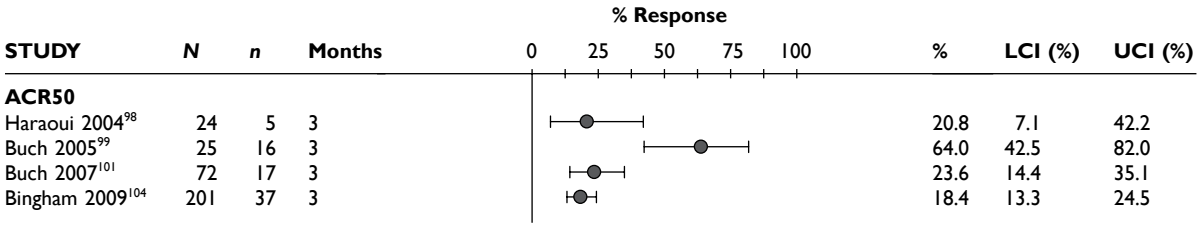
STUDY	N	n	Months	%	LCI (%)	UCI (%)
Any reason						
Haraoui 2004[98]	24	3	3	12.5	2.7	32.4
Bingham 2009[104]	201	13	3	6.5	3.5	10.8
Cohen 2005[100]	24	14	12	58.3	36.6	77.9
Laas 2009[103]	49	20	12	40.8	27.0	55.8
Adverse events						
Haraoui 2004[98]	24	0	3	0.0	0.0	14.2
Bingham 2009[104]	172	2	3	1.2	0.1	4.1
Cohen 2005[100]	24	3	12	12.5	2.7	32.4
Laas 2009[103]	49	8	12	16.3	7.3	29.7
Lack of efficacy						
Haraoui 2004[98]	24	0	3	0.0	0.0	14.2
Bingham 2009[104]	172	5	3	2.9	1.0	6.7
Cohen 2005[100]	24	7	12	29.2	12.6	51.1
Laas 2009[103]	49	12	12	24.5	13.3	38.9

FIGURE 7 Etanercept: withdrawals in the studies by reason. LCI, lower confidence interval; UCI, upper confidence interval.



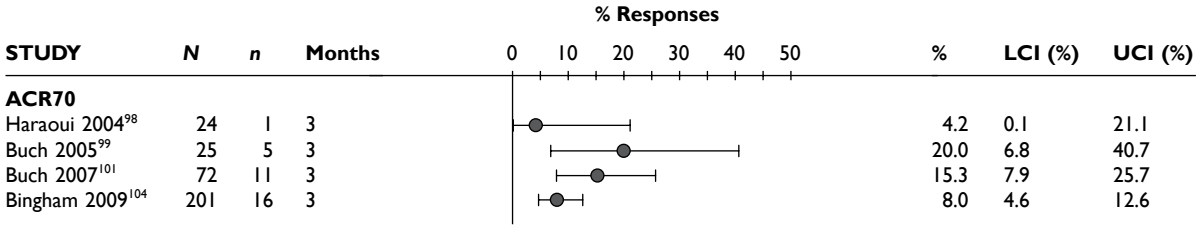
STUDY	N	n	Months	%	LCI (%)	UCI (%)
ACR20						
Haraoui 2004[98]	24	14	3	58.3	36.6	77.9
Buch 2005[99]	25	18	3	72.0	50.6	87.9
Buch 2007[101]	72	27	3	37.5	26.4	49.7
Bingham 2009[104]	201	85	3	42.3	35.4	49.4

FIGURE 8 ACR20: responses in patients receiving etanercept. LCI, lower confidence interval; UCI, upper confidence interval.



STUDY	N	n	Months	%	LCI (%)	UCI (%)
ACR50						
Haraoui 2004[98]	24	5	3	20.8	7.1	42.2
Buch 2005[99]	25	16	3	64.0	42.5	82.0
Buch 2007[101]	72	17	3	23.6	14.4	35.1
Bingham 2009[104]	201	37	3	18.4	13.3	24.5

FIGURE 9 ACR50: responses in patients receiving etanercept. LCI, lower confidence interval; UCI, upper confidence interval.



STUDY	N	n	Months	%	LCI (%)	UCI (%)
ACR70						
Haraoui 2004[98]	24	1	3	4.2	0.1	21.1
Buch 2005[99]	25	5	3	20.0	6.8	40.7
Buch 2007[101]	72	11	3	15.3	7.9	25.7
Bingham 2009[104]	201	16	3	8.0	4.6	12.6

FIGURE 10 ACR70: responses in patients receiving etanercept. LCI, lower confidence interval; UCI, upper confidence interval.

DAS

Figure 11 presents the mean changes from baseline in DAS. Four studies[100,101,103,104] reported using DAS28. The mean decrease in DAS28 ranged from 1.47 to 1.80 at 3 months. One study[102] reported no statistically significant decrease in DAS28 score from baseline at 12 months [mean change = –0.47, 95% confidence interval (CI) –1.06 to 0.12]. One study[102] reported DAS calculated based on 44 joints (DAS44). Iannone *et al.*[102] found no statistically significant differences in DAS44 scores when results for 16 and 24 weeks were compared with the baseline value.

EULAR response

Three studies reported EULAR responses. *Figure 12* shows the proportion of patients treated with ETN who achieved a good and good-to-moderate EULAR response after IFX failure. The percentage of patients who achieved a good score EULAR ranged from 12.5% to 45.8% at 3 months. The percentage of patients who achieved a good-to-moderate EULAR response ranged from 58.2% to 61.1% at 3 months.

Health Assessment Questionnaire

Three studies reported mean changes from baseline in HAQ score (*Figure 13*). In Haraoui *et al.*,[98] the change in HAQ score was –0.45. However, it was not reported whether this change was

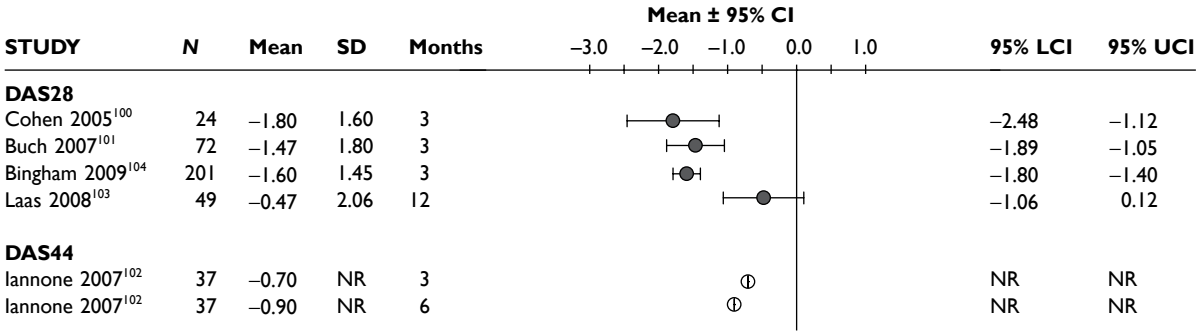


FIGURE 11 Etanercept: mean changes from baseline in DAS. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

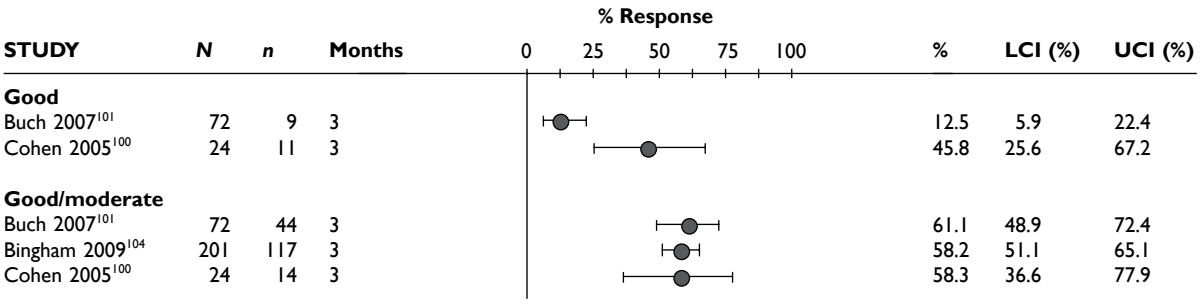


FIGURE 12 Etanercept: EULAR. LCI, lower confidence interval; UCI, upper confidence interval.

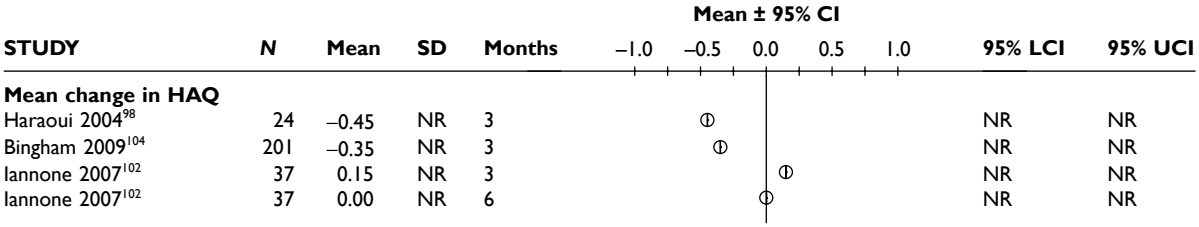


FIGURE 13 Etanercept: mean change from baseline in HAQ. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

statistically significant. For Iannone *et al.*,[102] the value of HAQ remained largely unchanged at 16 weeks (0.90) and 24 weeks (0.75) compared with the baseline value (0.75). In Bingham *et al.*,[104] there was a mean decrease in HAQ score of 0.35 at 3 months; this corresponds to a 22% decrease from baseline. This change was statistically significant.

One study[98] reported the percentage of patients who achieved minimal clinically important difference (MCID) in physical function (*Figure 14*). MCID was defined as a change of at least 0.22 in HAQ score. The percentage of patients who achieved MCID was 52%. Forty per cent of patients experienced an improvement in physical function of at least twice the value considered to represent MCID.

Quality of life

None of the studies assessed QoL.

Joint damage

None of the studies assessed joint damage.

Serious adverse events

Two studies reported serious AEs (*Figure 15*). Haraoui *et al.*[98] reported that no serious AEs occurred during the study. Bingham *et al.*[104] found that 5% of the patients experienced a serious AE during the study period.

Infection and serious infection

Three studies reported infection and serious infection (*Figure 16*). One study[104] reported that two patients (1%) experienced serious infections. The percentage of patients treated with ETN who reported any infection ranged from 4.1% to 8.3%.

Injection/infusion reaction

No study reported injection or infusion reaction.

Summary

For the assessment of effectiveness of ETN, seven uncontrolled studies were identified. Follow-up duration ranged from 12 weeks to over 9 months. Patients included in the studies were generally similar. The main results are summarised in *Table 14*.

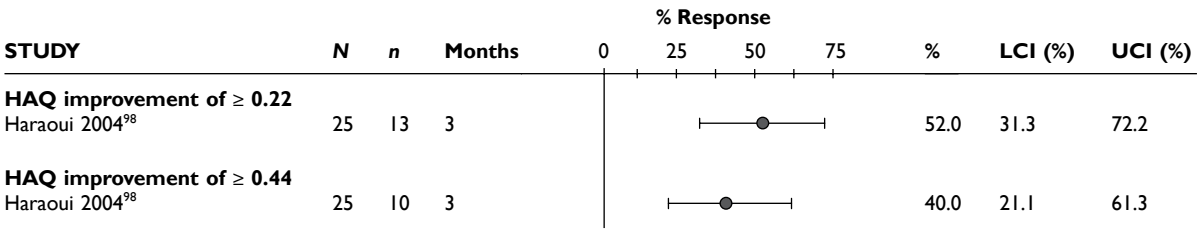


FIGURE 14 Etanercept: MCID physical function. LCI, lower confidence interval; UCI, upper confidence interval.

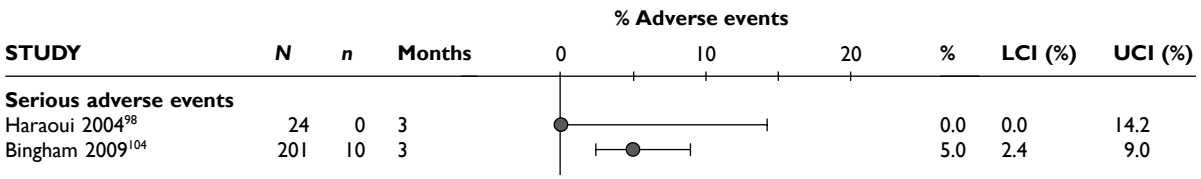


FIGURE 15 Etanercept: reported serious AEs. LCI, lower confidence interval; UCI, upper confidence interval.

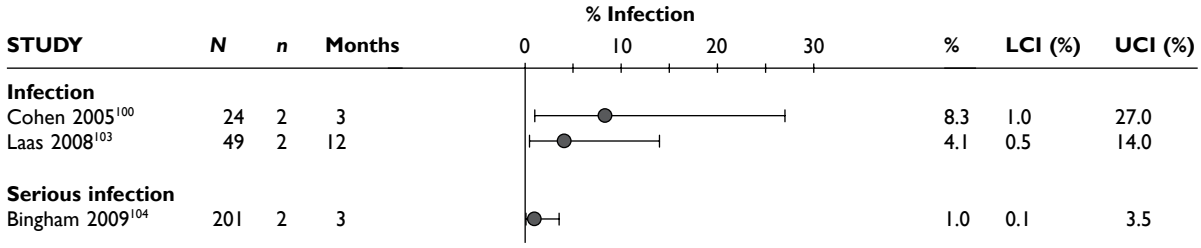


FIGURE 16 Etanercept: reported infection or serious infection. LCI, lower confidence interval; UCI, upper confidence interval.

TABLE 14 Etanercept: summary of main results

Outcome	3 months	6 months	≥9 months
Withdrawals (%):			
■ for any reason	6.5–12.5	NR	40.8–58.3
■ due to lack of efficacy	0.0–2.9	NR	24.5–29.2
■ due to AEs	0.0–1.2	NR	12.5–16.3
ACR20 response (%)	37.5–72.0	NR	NR
ACR50 response (%)	18.4–64.0	NR	NR
ACR70 response (%)	4.2–20.0	NR	NR
EULAR response (%):			
■ good/moderate response	58.2–61.1	NR	NR
■ good response	12.5–45.8	NR	NR
■ remission	NR	NR	NR
DAS28			
■ mean change from baseline	–1.47 to –1.80	NR	–0.47
DAS44			
■ mean change from baseline	–0.70	–0.90	NR
HAQ: mean change from baseline	–0.45 to 0.15	0.00	NR
QoL	NR	NR	NR
Joint damage	NR	NR	NR
Serious AEs (%)	0.0–5.0	NR	NR
Any infections/serious infections (%)	8.3/1.0	NR/NR	4.1/NR
Infusion reaction	NR	NR	NR

NR, not reported.

Infliximab

Overview of evidence

Three studies were identified that assessed IFX in comparison with standard care: one uncontrolled prospective study[107] and two uncontrolled retrospective studies.[105,106] (Note: the study by Yazici *et al.*[107] had a control group consisting of patients who were given their first biologic drug. This control group was not relevant to this report and, therefore, the study was utilised as uncontrolled.)

All included patients had tried one TNF inhibitor (ETN) before. Reasons for discontinuation included lack of efficacy, toxicity drug shortage, patient concerns about safety and thrombocytopenia.

All studies were conducted in the USA. Duration of follow-up was unclear in all the three studies.

Further details are provided in *Table 15*.

Patient characteristics

All three studies were rather small, with the number of patients treated with IFX ranging from 20 to 24. They provided very little information about the baseline characteristics of included patients. However, based on the available information there might have been some baseline differences between study populations (*Table 16*).

- In two studies the percentage of female participants ranged from 60% to 90%; Yazici *et al.*[107] did not provide any information.
- In two studies the mean age was 48 years and 61 years; it was not reported in Ang *et al.*[105]
- In two studies disease duration was 9.3 years and 13.4 years; it was not reported in Ang *et al.*[105]
- In two studies 34%–65% of patients were RF positive; no information was provided in Yazici *et al.*[107]
- In Ang *et al.*[105] 62% of patients were receiving MTX and 31% LEF; in Hansen *et al.*[106] all patients were receiving LEF and some of them also other DMARDs (AZA, sulfasalazine, MTX and prednisone); Yazici *et al.*[107] did not report concomitant DMARDs.
- Only one study (Hansen *et al.*[106]) reported that 75% of patients were receiving concomitant prednisone.
- Two studies reported the number of previous DMARDs – it ranged from 0 to over 5; it was not reported in Hansen *et al.*[106]
- Patients had tried one previous TNF inhibitor (ETN) in all three studies.
- None of the studies reported the baseline HAQ or DAS score.
- Only one study[106] reported that patients had a mean of 14 tender and 14 swollen joints at baseline.
- Only one study[106] reported the baseline ESR (mean 13 mm/hour) and CRP (mean 23.8 mg/dl).

TABLE 15 Infliximab: characteristics of included studies

Study	Country	Design	Reason for switching	Prior TNF inhibitors; *n*	Treatment arms (no. of patients)	Duration of follow-up	Comments
RCTs							
None were identified							
Non-randomised comparative studies							
None were identified							
Uncontrolled studies							
Ang 2003[105]	USA	Uncontrolled retrospective	Inadequate response, toxicity	ETN; 1	IFX (24)	Unclear	Average treatment duration 8.2 months
Hansen 2004[106]	USA	Uncontrolled retrospective	Lack of efficacy, drug shortage, patient concerns about safety, thrombocytopenia	ETN; 1	IFX (20)	Unclear	
Yazici 2004[107]	USA	Uncontrolled prospective	Inefficacy	ETN; 1	IFX (21); IFX (41)	Unclear	Group with 41 patients received IFX as first TNF inhibitor

DOI: 10.3310/hta15140

TABLE 16 Infliximab: baseline patient characteristics

Study	Number of patients/% female	Age (years), mean (SD)	RA duration (years), mean (SD)	RF positive (%)	% of patients on concomitant DMARDs and steroids	Number of previous DMARDs, mean (SD)	Number of previous TNF inhibitors, mean (SD)	HAQ, mean (SD)	DAS28, mean (SD)	TJC/SJC, mean (SD)	ESR (mm/hour), mean (SD)	CRP (mg/dl), mean (SD)
Ang 2003[105]	24/90[a]	NR	NR	34[a]	MTX (62), LEF (31)[a]	0 to > 5[a]	1	NR	NR	NR	NR	NR
Hansen 2004[106]	20/60	48 (NR)	9.3 (NR)	65	LEF (100); AZA (5); sulfasalazine (5); MTX (10); prednisone (75)	NR	1	NR	NR	14 (NR)/14 (NR)	13 (NR)	23.8 (NR)
Yazici 2004[107]	21/NR	61 (12)[b]	13.4 (9.8)[b]	NR	NR	2[b]	1	NR	NR	NR	NR	NR

NR, not reported; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count.
a The percentage is for data including the relevant study arm and an arm of five patients switched from IFX to ETN.
b Based on data for 88 patients including patients who were given IFX as the first biologic.

TABLE 17 Infliximab: non-RCT quality assessment

Study	Study design	Inclusion criteria clearly defined?	Were consecutive patients included in the study?	Patients withdrawn (%)	Comments
Ang 2003[105]	Uncontrolled retrospective	No	NA	Unclear	
Hansen 2004[106]	Uncontrolled retrospective	No	NA	Unclear	
Yazici 2004[107]	Uncontrolled prospective	No	Unclear	28.6	

NA, not applicable.

Quality assessment

Of the three identified studies, two were uncontrolled retrospective analyses. One study was uncontrolled and prospective. None of the studies reported inclusion criteria clearly. It was unclear if consecutive patients were included in Yazici *et al.*[107] and this item was not applicable to retrospective studies. A total of 28.6% were withdrawn from Yazici *et al.*[107] and this percentage was unclear in the remaining two studies. Details of the quality assessment are reported in *Table 7*.

Results

Table 18 indicates which of the outcomes reported in the main text of the report were assessed in individual studies and *Table 19* provides similar information for outcomes described in *Appendix 10* only.

Ang *et al.*[105] reported results in a way that made it impossible to utilise them in this report (correlations between response to IFX and ETN).

Withdrawals

Withdrawal for any reason was assessed only in Yazici *et al.*,[107] withdrawal because of lack of efficacy only in Hansen *et al.*[106] and withdrawal because of AEs was not assessed in any of the studies. Details are reported in *Figure 17*. Yazici *et al.*[107] reported that 28.6% of patients were withdrawn from the study for any reason (follow-up unclear). Ten per cent of patients were withdrawn from Hansen *et al.*[106] owing to lack of efficacy (follow-up unclear).

ACR20 response

None of the studies assessed ACR20 response.

ACR50 response

None of the studies assessed ACR50 response.

ACR70 response

None of the studies assessed ACR70 response.

DAS28

The only information on DAS28 change came from Yazici *et al.*[107] and the authors claimed that at 12 months patients 'improved significantly'.

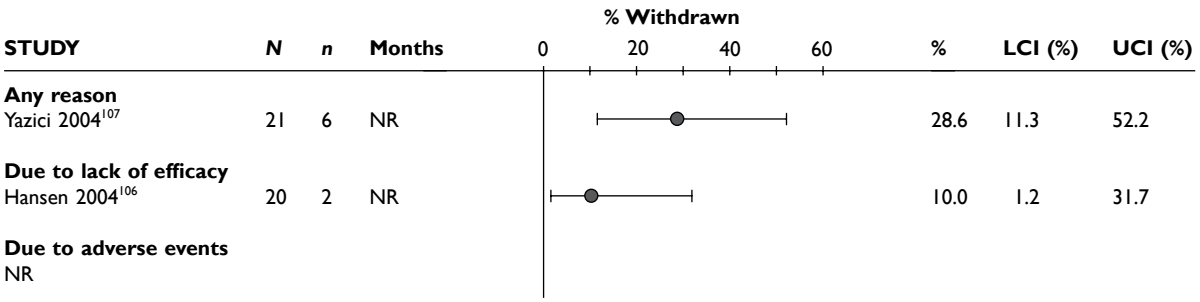


FIGURE 17 Infliximab: withdrawals. LCI, lower confidence interval; NR, not reported; UCI, upper confidence interval.

DOI: 10.3310/hta15140

TABLE 18 Infliximab: outcomes assessed in studies and reported in the main text of the report

Study	Total withdrawal	Withdrawal by reason	ACR (20/50/70)	DAS28	EULAR response	HAQ	QoL	Joint damage	Serious AEs	Infection/ serious infection	Injection/ infusion reaction
Ang 2003[105]											
Hansen 2004[106]		✓								✓	✓
Yazici 2004[107]	✓			✓							

TABLE 19 Infliximab: outcomes assessed in studies and reported in *Appendix 10* only

Study	Other measures of disease activity	Fatigue	Pain	TJC/SJC	CRP/ESR
Ang 2003[105]					
Hansen 2004[106]	✓			✓	✓
Yazici 2004[107]			✓		

SJC, swollen joint count; TJC, tender joint count.

EULAR response
The EULAR response was not assessed in any of the studies.

Health Assessment Questionnaire
The only information on HAQ change came from Yazici *et al.*[107] and the authors claimed that at 12 months patients 'improved significantly'.

Quality of life
Quality of life was not assessed in any of the studies.

Joint damage
Joint damage was not assessed in any of the studies.

Serious adverse events
Serious AEs were not assessed in any of the studies.

Infections/serious infections
Details of infections are reported in *Figure 18*. Fifteen percent of patients in Hansen *et al.*[106] experienced an infection (follow-up was unclear). No other studies reported infections. Serious infections were not reported in any of the studies.

Injection/infusion reaction
There were no infusion reactions in Hansen *et al.*[106] Other studies did not report this outcome.

Infliximab in comparison with an ongoing biologic agent
One RCT [open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept (OPPOSITE[133])] was identified that compared IFX with ongoing ETN. Although the study met the inclusion criteria of the systematic review, this comparison was not considered relevant to this report and, therefore, the study was not analysed.

It was a multicentre randomised trial and included 27 patients who had active RA and had an 'incomplete response to etanercept'. Patients were randomised either to discontinue ETN and receive IFX (13 patients) or to continue ETN treatment (14 patients). The follow-up duration was 30 weeks. Data were collected on outcomes including ACR response, HAQ, radiological progression, serum biomarker levels and safety.

Summary
Three studies compared IFX with standard care: two uncontrolled retrospectiv[105,106] and one uncontrolled prospective Yazici *et al.* studies.[107] They included small numbers of patients ranging from 20 to 24. Follow-up was unclear in all of them. There was little information about baseline characteristics; however, it seems that there may be some, if small, differences between studies. The main results of included studies are summarised in *Table 20*.

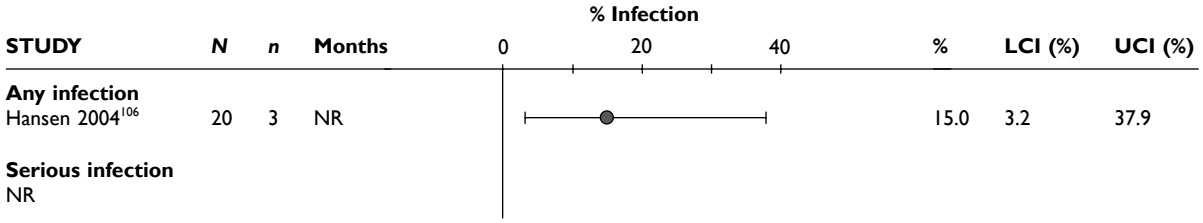


FIGURE 18 Infliximab: infections. LCI, lower confidence interval; NR, not reported; UCI, upper confidence interval.

TABLE 20 Infliximab: summary of main results

	Uncontrolled studies
Outcome	**Unclear follow-up**
Withdrawals (%):	
■ for any reason	28.6 (reported in one study)
■ due to lack of efficacy	10 (reported in one study)
■ due to AEs	NR
ACR20 response	NR
ACR50 response	NR
ACR70 response	NR
DAS28	Only one study included a statement that at 12 months patients 'improved significantly'
EULAR response	NR
HAQ	Only one study included a statement that at 12 months patients 'improved significantly'
QoL	NR
Joint damage	NR
Serious AEs	NR
Any infections/serious infections (%)	15 (reported in one study)/NR
Infusion reaction	0 (reported in one study)

NR, not reported.

Tumour necrosis factor inhibitors as a class

Overview of evidence

This section reports on studies that evaluated the use of TNF inhibitors as a class after the failure of the first one. No RCT was found. One controlled[121–123] and six uncontrolled observational studies[108–113] were identified. In Finckh *et al.*[136,137] lack of efficacy was the primary reason for switching TNF inhibitors. In Hyrich *et al.*,[121–123] Gomez-Reino *et al.*[108] and Blom *et al.*[113] patients switched to another TNF inhibitor because of a lack of efficacy or AEs. In Hjardem *et al.*,[110] Duftner *et al.*[111] and Karlsson *et al.*[112] patients switched TNF inhibitorsowing to lack of efficacy or AEs or for other reasons. The reason for changing from one TNF inhibitor to another was unclear in Solau-Gervais *et al.*[109] Hyrich *et al.*[121–123] used data from the BSRBR. The other studies were carried out in Switzerland, Spain, France, Denmark, Austria, Sweden and the Netherlands. The length of follow-up ranged from 3 months to up to 4 years. Further details are provided in *Table 21*.

TABLE 21 Tumour necrosis factor inhibitors as class: characteristics of included studies

Study	Country	Design	Reason for switching	Prior TNF inhibitors (*n*)	Treatment arms (no. of patients)	Duration of follow-up	Comments
RCTs							
None were identified							
Non-randomised controlled studies							
Hyrich 2009[121–123]	UK	Cohort	Inefficacy, AEs	ETN, IFX, ADA	TNF inhibitor (all switchers: *n* = 534; stoppers: *n* = 202)	> 6 months	
Uncontrolled studies							
Gomez-Reino 2006[108]	Spain	Uncontrolled prospective	AEs, lack of efficacy	IFX, ETN	TNF inhibitor (*n* = 448)	2 years	Including other forms of arthritis (ankylosing spondylitis, psoriatic arthritis and other chronic arthritis; *n* = 385 for RA)
Solau-Gervais 2006[109]	France	Uncontrolled prospective	Unclear	Any	TNF inhibitor (*n* = 70)	> 3 months	
Hjardem 2007[110]	Denmark	Uncontrolled retrospective	Inefficacy, AE, other	ETN, IFX, ADA	TNF inhibitor (*n* = 235)	3 months	
Duftner 2008[111]	Austria	Uncontrolled retrospective	Inefficacy, AE, other	IFX, ETN, ADA	TNF inhibitor (*n* = 109)	< 4 years	Length of follow-up including first line; reported 12-month drug continuation rate for second, third and fourth line
Karlsson 2008[112]	Sweden	Uncontrolled retrospective	Inefficacy, AE, other	Any	TNF inhibitor (*n* = 337)	3 months	Second and third line separately
Blom 2009[113]	Netherlands	Uncontrolled retrospective	Non-response, loss of response, and AEs	IFX, ETN, ADA	IFX, ETN, ADA (*n* = 197)	6 months	
Finckh 2009[136,137]	Switzerland	Prospective cohort	Inadequate response	Any (≥ 1)	RTX (*n* = 155); alternative TNF inhibitor (*n* = 163)	11 months (median)	Based on the Swiss Clinical Quality Management program for Rheumatoid Arthritis (SCQM-RA)

Patient characteristics

Full details of patients' characteristics are reported in *Table 22*. The number of patients included in the studies ranged from 70 to 818. Patient characteristics were generally similar across the eight studies:

- The percentage of female patients ranged from 67% to 89%.
- Where reported, the mean age ranged from 51 years to 58 years.
- Where reported, the mean disease duration ranged from 8.0 years to 14.7 years.
- Where reported, the percentage of RF-positive patients ranged from 51.5% to 81%.
- Where reported, 61%–75% patients were on MTX; 55%–68% of patients were receiving corticosteroids.
- Where reported, the mean number of previously used conventional DMARDs varied from 4.0 to 4.7.
- Where reported, studies included patients who previously tried IFX, ETN and ADA.
- Where reported, the mean baseline HAQ ranged from 1.4 to 1.9.

TABLE 22 Tumour necrosis factor inhibitors as a class: baseline patient characteristics

Study	Number of patients/% female	Age (years), mean (SD)	RA duration (years), mean (SD)	RF positive (%)	% of patients on concomitant DMARDs and steroids	Number of previous DMARDs, mean (SD)	Number of previous TNF inhibitors, mean (SD)	HAQ, mean (SD)	DAS28, mean (SD)	TJC/SJC, mean (SD)	ESR (mm/hour), mean (SD)	CRP (mm/dl), mean (SD)
Hyrich 2009[121–123]	818/80	58 (11)	10.0 (8.9)	NR	NR	4.0 (1.5)	NR	1.90 (0.63)	6.5 (1.0)	NR	NR	NR
Gomez-Reino 2006[108]	448/67	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Solau-Gervais 2006[109]	70/86	57 (NR)	14.7 (NR)	81	NR	4.3 (NR)	≥1	NR	6.0 (NR)	NR	NR	NR
Hjardem 2007[110]	235/75	55 (12)	8.0 (9.6)	NR	MTX (75), prednisone, corticosteroids	4.0 (1.5)	≥1	NR	5.2 (1.3)	NR	NR	NR
Duftner 2008[111]	109/89	51 (12)	8.0 (7.5)	NR	NR	NR	≥1	NR	NR	NR	NR	NR
Karlsson 2008[112]	337/82	56 (13)	14.0 (10.0)	NR	MTX, corticosteroid (68)	4.7 (1.9)	1–3	1.40 (0.60)	5.5 (1.3)	9.3 (6.8)/8.4 (5.9)	36 (25)	28 (35)
Blom 2009[113]	197/71	55 (NR)	7.9 (NR)	52	MTX, steroids	NR	NR	NR	5.1 (1.2)	NR	NR	NR
Finckh 2009[136,137]	163/78	55 (13)	11.0 (7.0)	77	MTX 61, steroids 55	NR	1	1.42 (0.96–1.85)[a]	4.1 (1.3)	NR	NR	NR

NR, not reported; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count.
a Median and interquartile range.

TABLE 23 Tumour necrosis factor inhibitors as class: non-RCT quality assessment

Study	Design	Inclusion criteria clearly defined?	Were consecutive patients included in the study?	Patients withdrawn (%)	Comments
Hyrich 2009[121–123]	Controlled prospective	Yes	Unclear	Unclear	
Gomez-Reino 2006[108]	Uncontrolled prospective	Unclear	Unclear	Unclear	
Solau-Gervais 2006[109]	Uncontrolled retrospective	Yes	NA	Unclear	
Hjardem 2007[110]	Uncontrolled retrospective	Unclear	Unclear	34.5	
Duftner 2008[111]	Uncontrolled retrospective	Yes	NA	NA	
Karlsson 2008[112]	Uncontrolled prospective	Yes	No	NA	140/477 excluded (see text)
Blom 2009[113]	Uncontrolled retrospective	Yes	NA	38.6	
Finckh 2009[136,137]	Prospective cohort	Yes	Unclear	NA	

NA, not applicable.

- Where reported, the mean DAS28 score ranged from 4.1 to 6.5.
- The mean number of tender and swollen joints was reported only in one study (tender 9.3 and swollen 8.4).
- The mean baseline ESR was reported in one study and was 36 mm/hour.
- The baseline CRP was reported in one study and was 2.8 mg/dl.

Quality assessment

One study was controlled.[121–123] Two studies[108,109] were uncontrolled and prospective. Four studies were uncontrolled and retrospective.[110–113] Finckh *et al.*[136,138] was a non-randomised comparative study (TNF inhibitors vs RTX). This section presents data only for TNF inhibitors. Full details of the quality assessment are reported in *Table 23*. Most studies stated clearly their inclusion criteria. The inclusion criteria were unclear in two studies.[108,110] It was unclear in most studies whether consecutive patients were included in the study. Nearly one-third (140/477) of patients who met the study inclusion criteria were excluded from Karlsson *et al.*[112] because of dropouts/missing response data at 3 months. The exclusion of these patients may partly account for the higher rates of EULAR responses observed in this study compared with other studies (see *Figure 25*). The percentage of patients withdrawn was clearly reported in two studies.

Results

Tables 24 and *25* state which outcomes were measured in each study and whether they are reported in the main text or *Appendix 10* of this report.

Withdrawals

Two studies reported withdrawals together with reasons for withdrawing treatment (*Figure 19*). The percentage of patients who withdrew for any reason ranged from 7.6% (at 3 months) to 38.6% (at 12 months). The percentage of patients who withdrew because of AEs ranged from 6.1% (at 3 months) to 10.2% (at 6 months). At 12 months, the percentage of patients who withdrew because of AEs ranged from 6.0% to 14.7%. The percentage of patients who withdrew because of lack of efficacy ranged from 1.5% (at 3 months) to 22.6% (at 12 months).

One study reported 1-year drug survival[108] (probability of staying on treatment at 12 months) of 0.79 (95% CI 0.74 to 0.83). Two studies reported median drug survival.[110,111] Hjardem *et al.*[110] and Duftner *et al.*[111] reported that the median drug survival was 37 weeks and 8.0 months (range 0–43.7 months), respectively.

ACR20 response

The ACR20 response was assessed in one study (*Figure 20*). Karlsson *et al.*[112] reported that at 3 months ACR20 response rate was 49.0% (95% CI 43.5% to 54.4%).

ACR50 response

The ACR50 response was assessed in one study (*Figure 21*). Karlsson *et al.*[112] reported that at 3 months ACR50 response rate was 25.8% (95% CI 21.2% to 30.8%).

ACR70 response

The ACR70 response was assessed in one study (*Figure 22*). Karlsson *et al.*[112] reported that at 3 months ACR70 response rate was 7.1% (95% CI 4.6% to 10.4%).

DAS28

Three studies reported mean changes from baseline in the DAS28 score (*Figure 23*). The mean decrease in DAS28 ranged from 0.86 to 1.00 at 3 months and from 0.88 to 0.92 at 6 months. Two studies[112,113] reported low disease activity (DAS28 less than 3.2) (*Figure 24*). At 3 months the

DOI: 10.3310/hta15140

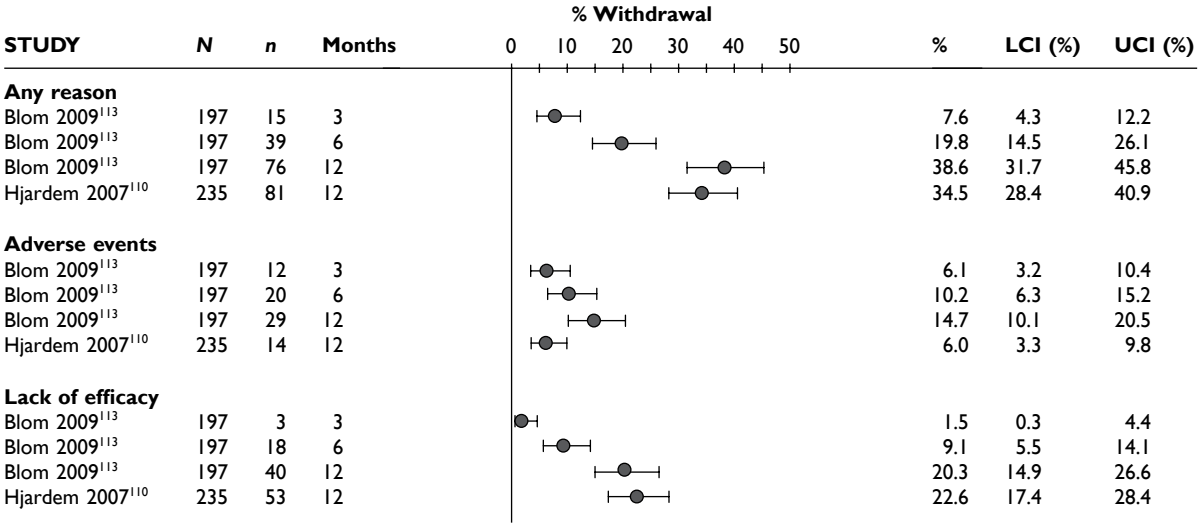
TABLE 24 Tumour necrosis factor inhibitors as class: outcomes assessed in studies and reported in the main text of the report

Study	Total withdrawal	Withdrawal by reason	ACR (20/50/70)	DAS28	EULAR response	HAQ	QoL	Joint damage	Serious AEs	Infection/ serious infection	Injection/ infusion reaction
Hyrich 2009[121–123]						✓					
Gomez-Reino 2006[108]											
Solau-Gervais 2006[109]											
Hjardem 2007[110]	✓	✓		✓	✓				✓	✓	
Duftner 2008[111]	✓									✓	
Karlsson 2008[112]			✓	✓	✓						
Blom 2009[113]	✓	✓		✓	✓						
Finckh 2009[136,137]				✓							

TABLE 25 Tumour necrosis factor inhibitors as class: outcomes assessed in studies and reported in the *Appendix 10* only

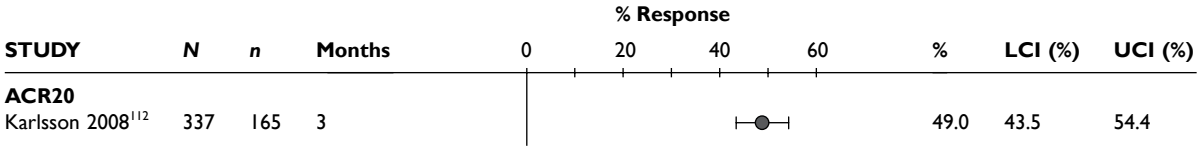
Study	Other measures of disease activity	Fatigue	Pain	TJC/SJC	CRP/ESR
Hyrich 2009[121–123]					
Gomez-Reino 2006[108]					
Solau-Gervais 2006[109]					
Hjardem 2007[110]					
Duftner 2008[111]					
Karlsson 2008[112]					
Blom 2009[113]					
Finckh 2009[136,137]				✓	✓

SJC, swollen joint count; TJC, tender joint count.



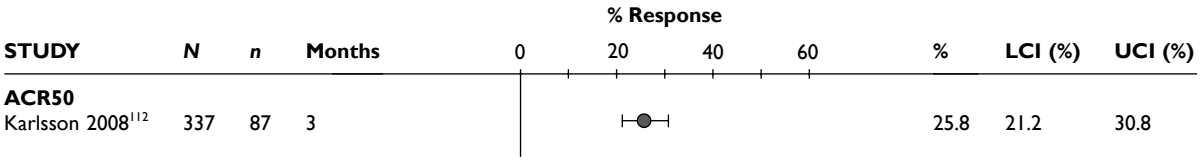
STUDY	N	n	Months	%	LCI (%)	UCI (%)
Any reason						
Blom 2009[113]	197	15	3	7.6	4.3	12.2
Blom 2009[113]	197	39	6	19.8	14.5	26.1
Blom 2009[113]	197	76	12	38.6	31.7	45.8
Hjardem 2007[110]	235	81	12	34.5	28.4	40.9
Adverse events						
Blom 2009[113]	197	12	3	6.1	3.2	10.4
Blom 2009[113]	197	20	6	10.2	6.3	15.2
Blom 2009[113]	197	29	12	14.7	10.1	20.5
Hjardem 2007[110]	235	14	12	6.0	3.3	9.8
Lack of efficacy						
Blom 2009[113]	197	3	3	1.5	0.3	4.4
Blom 2009[113]	197	18	6	9.1	5.5	14.1
Blom 2009[113]	197	40	12	20.3	14.9	26.6
Hjardem 2007[110]	235	53	12	22.6	17.4	28.4

FIGURE 19 Tumour necrosis factor inhibitor as a class: withdrawals from the studies by reason. LCI, lower confidence interval; UCI, upper confidence interval.



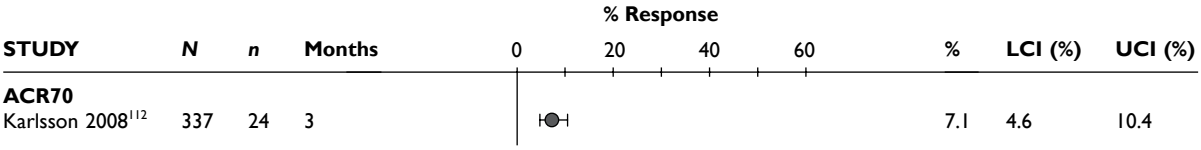
STUDY	N	n	Months	%	LCI (%)	UCI (%)
ACR20						
Karlsson 2008[112]	337	165	3	49.0	43.5	54.4

FIGURE 20 TNF inhibitors as a class: ACR20 response. LCI, lower confidence interval; UCI, upper confidence interval.



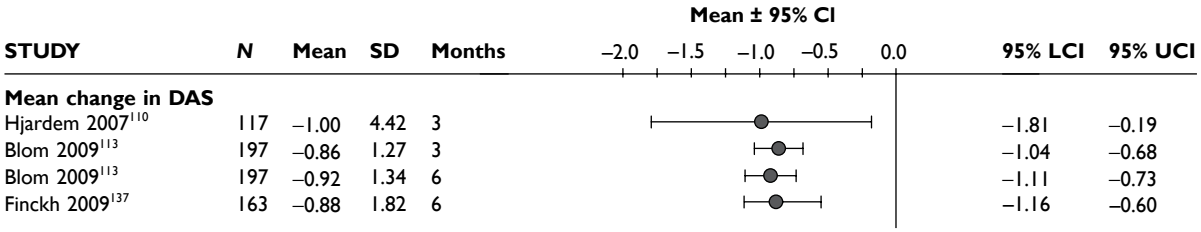
STUDY	N	n	Months	%	LCI (%)	UCI (%)
ACR50						
Karlsson 2008[112]	337	87	3	25.8	21.2	30.8

FIGURE 21 TNF inhibitors as a class: ACR 50 response. LCI, lower confidence interval; UCI, upper confidence interval.



STUDY	N	n	Months	%	LCI (%)	UCI (%)
ACR70						
Karlsson 2008[112]	337	24	3	7.1	4.6	10.4

FIGURE 22 TNF inhibitor as a class: ACR70 response. LCI, lower confidence interval; UCI, upper confidence interval.



STUDY	N	Mean	SD	Months	95% LCI	95% UCI
Mean change in DAS						
Hjardem 2007[110]	117	–1.00	4.42	3	–1.81	–0.19
Blom 2009[113]	197	–0.86	1.27	3	–1.04	–0.68
Blom 2009[113]	197	–0.92	1.34	6	–1.11	–0.73
Finckh 2009[137]	163	–0.88	1.82	6	–1.16	–0.60

FIGURE 23 TNF inhibitors as a class: mean changes from baseline in DAS28. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

percentage of patients with low disease activity ranged from 14.2% to 29.1%. One study reported DAS28 remission (DAS28 less than 2.6) (*Figure 24*). Karlsson *et al.*[112] reported that 15.4% (95% CI 11.7% to 19.7%) of patients were in remission.

EULAR response

Three studies[110,112,113] reported the percentage of patients who achieved good and good-to-moderate EULAR responses (*Figure 25*). The percentage of patients who achieved good EULAR response ranged from 8.6% to 22.8% at 3 months and was 9.1% at 6 months. The percentage of patients who achieved good-to-moderate EULAR response ranged from 31.5% to 64.7% at 3 months. Only one study reported good-to-moderate EULAR response at 6 months (32.5%).

Health Assessment Questionnaire

Only one study reported mean changes from baseline in HAQ score (*Figure 26*). Hyrich *et al.*[121–123] compared patients who discontinued TNF inhibitor within the first 12 months and did not start a subsequent TNF inhibitor or other biologic drug during the next 12 months ('stoppers') with patients who stopped their first TNF inhibitor within the first 12 months of therapy because of the lack of efficacy, but started a second TNF inhibitor during the subsequent 12 months ('switchers'). The mean change in HAQ score was adjusted for differences in age, gender, disease duration, HAQ score at first failure, DAS28 at start of first TNF inhibitor and DAS28 score at first failure. 'Switchers' (adjusted mean change = –0.11, 95% CI –0.18 to –0.04) had significantly greater improvement in HAQ score than 'stoppers' (*Figure 26*).

Quality of life

None of the studies reported QoL.

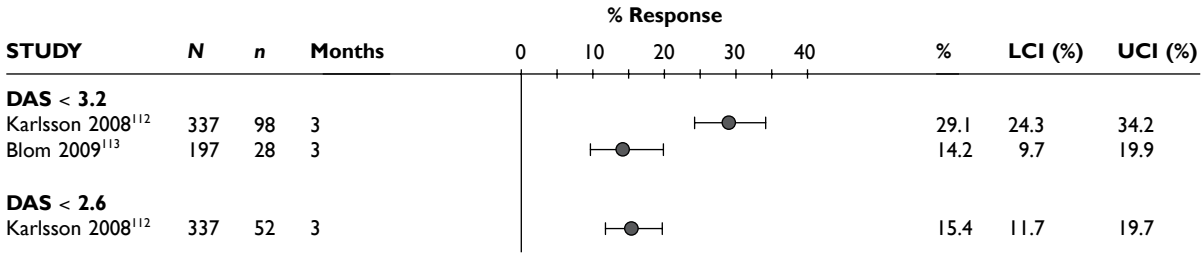


FIGURE 24 TNF inhibitors as a class: low disease activity (DAS28 <3.2) and remission (DAS28 <2.6). LCI, lower confidence interval; UCI, upper confidence interval.

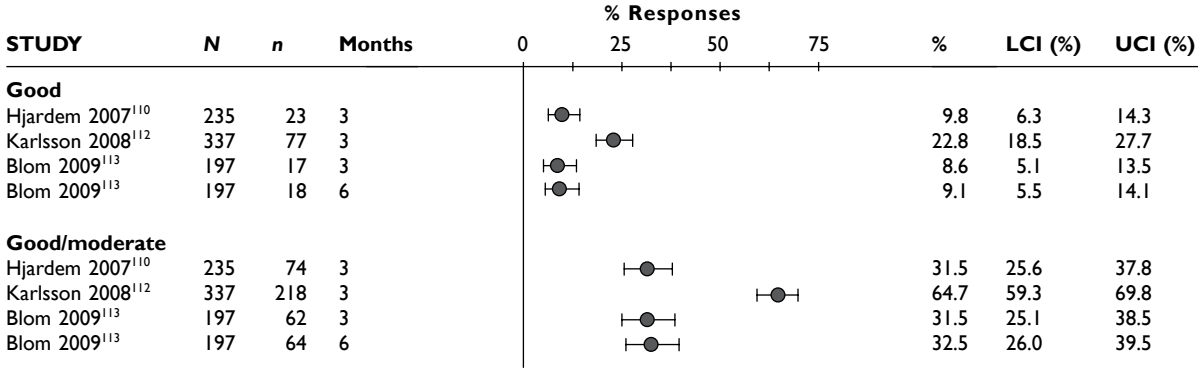


FIGURE 25 TNF inhibitors as a class: EULAR response rates. LCI, lower confidence interval; UCI, upper confidence interval.

Joint damage

None of the studies reported joint damage.

Serious adverse events

Only one study reported serious AEs (*Figure 27*). Hjardem *et al.*[110] reported that 6.0% (95% CI 3.3% to 9.8%) of the patients experienced a serious AE during the study period.

Infection and serious infection

Two studies reported infection and serious infection (*Figure 28*). At 3 months the percentage of patients who experienced infection ranged from 27.2% to 28.1%. One study[111] reported that 13.9% (95% CI 9.1% to 19.9%) of the patients experienced serious infections at 3 months.

Injection/infusion reaction

None of the studies reported injection or infusion reactions.

Summary

For the assessment of effectiveness of TNF inhibitors as a class after failure of the first TNF inhibitor, one non-randomised comparative and seven uncontrolled studies were identified. Follow-up duration ranged from 3 months to 4 years. Patients included in the studies were generally similar. The main results are summarised in *Table 26*.

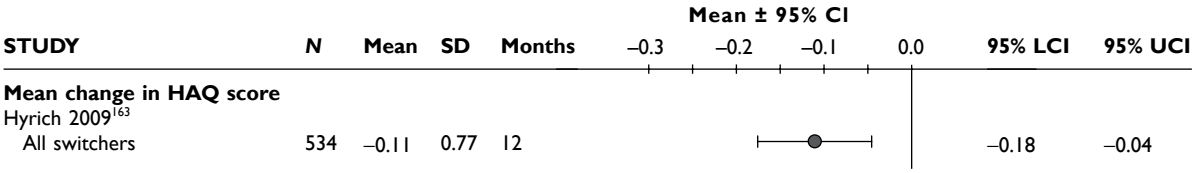


FIGURE 26 TNF inhibitors as a class: adjusted mean change from baseline in HAQ score. Adjusted for age, gender, disease duration, HAQ score at first failure, DAS28 at start of first TNF inhibitor and DAS28 score at first failure. LCI, lower confidence interval; UCI, upper confidence interval.

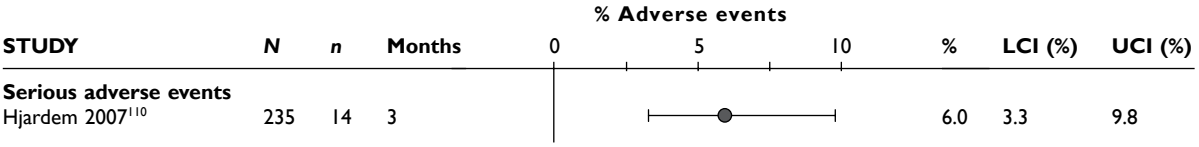


FIGURE 27 TNF inhibitors as a class: serious adverse events. LCI, lower confidence interval; UCI, upper confidence interval.

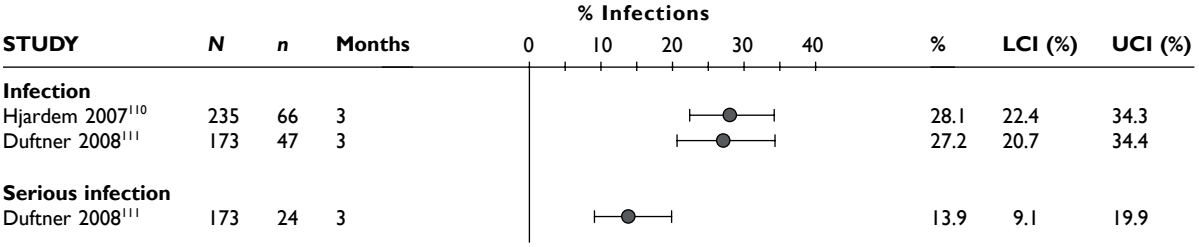


FIGURE 28 TNF inhibitors as a class: infections and serious infections. LCI, lower confidence interval; UCI, upper confidence interval.

TABLE 26 TNF inhibitors as a class: summary of main results

Outcome	3 months	6 months	≥9 months
Withdrawals (%):			
■ for any reason	7.6	19.8	34.5–38.6
■ due to lack of efficacy	1.5	9.1	20.3–22.6
■ due to AEs	6.1	10.2	6.0–14.7
ACR20 response (%)	49.0	NR	NR
ACR50 response (%)	25.8	NR	NR
ACR70 response (%)	7.1	NR	NR
EULAR response (%):			
■ good/moderate response	31.5–64.7	32.5	NR
■ good response	8.6–22.8	9.1	NR
■ remission	NR	NR	NR
DAS28: mean change from baseline			
	–1.00 to –0.86	–0.92 to –0.88	NR
DAS28 < 3.2 (%)	14.2–29.1	NR	NR
DAS28 < 2.6 (%)	15.4	NR	NR
HAQ: mean change from baseline	NR	NR	–0.11[a]
QoL	NR	NR	NR
Joint damage	NR	NR	NR
Serious AEs	6.0%	NR	NR
Any infections/serious infections (%)	27.2–28.1/13.9	NR	NR
		NR	NR
Infusion reaction	NR	NR	NR

NR, not reported.
a Adjusted for age, gender, disease duration, HAQ score at first failure, DAS28 at start of first TNF inhibitor and DAS28 score at first failure.

Rituximab

Overview of evidence

Seven studies were identified that assessed RTX: one RCT [randomised evaluation of long-term efficacy of rituximab in rheumatoid arthritis (REFLEX)[124–126]] and six uncontrolled studies.[114–118,137,139] One of these (Finckh *et al.*[136,137]) contained a comparative arm with an alternative TNF inhibitor; the comparative data are described in the section *Evidence from comparative studies*. One study[116] included data from patients of whom nearly half were previously TNF inhibitor naive. Only data reported separately for those who had a previous TNF inhibitor were included in this report. In another study,[118] at 6 months, 17 patients (including five who were TNF inhibitor naive at original baseline) started a second course of TNF inhibitor; data for this group of patients were excluded from the report.

Data from one cohort analysis of the REFLEX RCT extension[139] and one pooled analysis of all RTX development studies from the MS are also described. The REFLEX extension[139] was a long-term follow-up analysis of repeated treatment data of the original RCT: it included patients who had responded to an initial course of RTX during the RCT and received open-label treatment with the same RTX regimen for up to three repeat treatment courses. (Note: responding patients in the initial REFLEX RCT[124–126] after reaching the primary end point at week 24 requiring further courses of RTX treatment entered the open-extension study.) Patients from the placebo arm of the RCT were also included and received their first course of RTX within the extension

study. A total of 480 patients from the RCT (308 from the RTX arm and 172 from the placebo arm) entered the extension phase.

The manufacturer's pooled analysis combined data from patients of the REFLEX RCT,[124–126] together with data from its open-label extension study, and from other studies in manufacturer's RTX development programme. (Note: data were pooled for patients who only received the expected licensed dose of RTX two × 1,000 mg plus MTX regimen for first and subsequent courses and who received prior TNF inhibitor therapy.) It is unclear how many patients from the REFLEX trial[124–126] were included in the pooled analysis.

The Keystone *et al.* uncontrolled study[116] also reported data for up to two treatment courses; these data are presented with those from the REFLEX extension[139] and the RTX pooled analysis.

The REFLEX trial was a multicentre RCT conducted in 114 counties in the USA, Europe, Canada and Israel. Of the six uncontrolled studies, one was conducted in Switzerland, one in the UK, one in Sweden, one in the Netherlands and one in France. For the studies included in the Keystone *et al.* analysis,[116] and for those included in the manufacturer's pooled analysis, except the REFLEX trial,[139] it is unclear in which country the studies were conducted.

Further details are provided in *Table 27*.

Patient characteristics

Data on patient baseline characteristics can be found in *Table 28*. Patient characteristics were not reported for the manufacturer's pooled analysis and were not reported separately for the patients who had previously received a TNF inhibitor in the Keystone *et al.* analysis.[116]

The number of patients included in the REFLEX RCT[124–126] was 517 and ranged from 20 to 155 in the six uncontrolled studies. Where reported, characteristics of the patients included in the studies varied in some aspects, but were generally similar:

- The percentage of female patients ranged from 77% to 86%.
- The mean age ranged from 52 to 58 years in four studies and the median age in two studies was 54–55 years.
- The mean disease duration ranged from 10 to 15 years in four studies and the median age in two studies was 12–16 years.
- The percentage of RF-positive patients ranged from 79% to 90% and was lowest in the REFELX study; one study and both analyses from the MS did not report this.
- Concomitant DMARDs were reported in five studies: 30%–100% patients were on MTX; all the patients in the REFLEX RCT[124–126] were on concomitant DMARDs.
- The proportion of patients who were receiving concurrent steroids ranged from 55% to 100%; one study did not report this.
- The mean number of previously used conventional DMARDs reported in three studies ranged from 2.5 to 4.2 and median reported in the other two ranged from 3 to 4.
- Where reported, the mean number of previous TNF inhibitors was 1 or greater than 1 and the median number reported in two uncontrolled studies was 2.
- The mean baseline HAQ was reported only in the REFELX study was 1.9 and the median baseline HAQ reported in two uncontrolled studies ranged from 1.6 to 2.6.
- Where reported, the mean DAS28 score ranged from 5.0 to 6.9 and it was the highest in the REFELX study.
- The mean number of tender joints was 34 and swollen joints was 23 in the REFLEX trial; [124–126] the median number was 26 and 13 respectively in Jois *et al.*;[115] other studies did not report the baseline number of tender and swollen joints.

TABLE 27 Rituximab: characteristics of included studies

Study	Country	Design	Reason for switching	Prior anti-TNFs (no.)	Treatment arms (*n* of patients)	Duration of follow-up	Comments
RCTs							
REFLEX[124–126]	North America, Europe, Israel	Prospective randomised controlled parallel	Inadequate response or intolerance	Any (≥1)	RTX (*n*=308) PL (*n*=209)	24 weeks; 48 weeks[a]	Pivotal trial for anti-TNF inadequate responders
Uncontrolled studies							
Bokarewa 2007[114]	Sweden	Prospective uncontrolled	Lack of response	Any biologic (*n* unclear)	RTX (*n*=48)	12 months	Dosing schedule different from licence; not only TNF inhibitor failures; a few patients tried other biologics (anti-thymocyte globulin treatment, IL-1 receptor antagonist); 64% had experienced more than one biologic drug prior to RTX treatment
Jois 2007[115]	UK	Prospective uncontrolled	Lack of response	Any (≥2)	RTX (*n*=20)	6 months	All patients had failed at least two TNF inhibitors (10 had failed three TNF inhibitors, five also failed anakinra) Patients were offered retreatment with a second cycle of RTX if they had responded to the earlier one but flared
Keystone 2007[116]	Unclear	Retrospective uncontrolled	Unclear	All had TNF inhibitor (*n* unclear)	RTX (*n*=155 to 158[b])	6 months[b]	A pooled analysis of 1,039 patients who received ≥ 1 courses of RTX, 427 (41%) of whom were previously TNF inhibitor naive. 570 of these patients had ≥ 2 courses of RTX, 255 (45%) of whom were previously TNF inhibitor naive. Only data that were reported separately for those who had prior TNF inhibitor were included in this report
Assous 2008[117,c]	France	Retrospective uncontrolled	Lack of response; contraindication	Any (*n* unclear)	RTX (*n*=50)	6 months	20/50 patients had contraindications to TNF inhibitors; previous exposure to TNF inhibitor treatment was not clear in these patients
Thurlings 2008[118]	Netherlands	Prospective uncontrolled	Side effects; inefficacy	Any (=1; >1?)	RTX (*n*=30)	6 months	Five patients were TNF inhibitor naive; at 6 months 17 patients including the five who were TNF inhibitor naive were retreated with a second RTX course, seven patients (unclear how many of them were TNF inhibitor naive at the beginning of the study) were retreated later with a third RTX course
Finckh 2009[136,137]	Switzerland	Prospective cohort	Inadequate response	Any (≥1)	RTX (*n*=155)	11 months (median)	Based on the SCQM-RA
REFLEX extension[139]	North America, Europe, Israel	Uncontrolled retrospective	Inadequate response or intolerance	Any (≥1)	RTX (*n*=480)		308 were from the RTX arm and 172 were from the placebo arm. Of these, 307 received two courses, 235 received three courses, 146 received four courses and 58 received five courses
RTX pooled analysis[139,d]	NA	Uncontrolled retrospective	NR	Any (≥1)	RTX		A pooled analysis of data derived primarily from REFLEX RCT[124–126] and its extension study,[139] and also other studies in Roche's RTX development programme. Data included only for patients who received previous TNF inhibitor treatment and who had received the licensed dose of RTX for first and subsequent courses

NA, not applicable; NR, not reported; SCQM–RA, Swiss-RA, Swiss Clinical Quality Management Program for Rheumatoid Arthritis.

a REFLEX – data of long-term efficacy from a single course of RTX from the Roche submission.[137]

b Evaluable patients 24 weeks after two courses of RTX treatment and who previously received a prior TNF inhibitor.

c Only 30/50 patients who had inadequate response to previous anti-TNF treatment were included in the main analysis in this report. The remaining 20 patients were included only in sensitivity analyses.

d Data from the Roche submission.[137] Included were patients from the primary REFLEX study,[122–124] its open-label extension study, and other studies in the RTX development programme, data have been pooled for patients who only received the expected licensed dose of RTX two × 1,000 mg + MTX regimen for first and subsequent courses and who received prior TNF inhibitor therapy.

Bold type indicates that the study was an RCT.

DOI: 10.3310/hta15140

TABLE 28 Rituximab: baseline patient characteristics

Study	Number of patients/% female	Age (years), mean (SD)	RA duration (years), mean (SD)	RF positive (%)	% of patients on concomitant DMARDs and steroids	Number of previous DMARDs, mean (SD)	Number of previous TNF inhibitors, mean (SD)	HAQ, mean (SD)	DAS28, mean (SD)	TJC/SJC, mean (SD)	ESR (mm/hour), mean (SD)	CRP (mg/dl), mean (SD)
REFLEX[124–126]	517/81	52 (12)	12 (8)	79	MTX 100; steroids 63	2.5 (1.8)	1.5 (0.7)	1.90 (0.60)	6.9 (1.0)	34 (15)/23 (12)	48 (27)	3.7 (3.9)
Bokarewa 2007[114]	48/79	58 (11)	10 (7)	83	MTX 77; steroids 42	4.2 (range 3.0–8.0)	64% had had >1 biologic previously	NR	6.1 mean (range 4.0 to 7.8)	NR[b]	NR[b]	NR[b]
Jois 2007[115]	20/80	54 (33–80)[d]	16 (5–39)[d]	90	MTX 30; steroids 60	3.0 (2.0–8.0)[d]	2 (2–4)[b]	2.60 (0.75–3.00)[d]	7.2 (5.3–9.0)[d]	26 (2–28)[d]/13 (0–26)[d]	56 (14–125)[d]	3.2 (0.3–17.4)
Keystone 2007[116,e]	NR	NR	NR	NR	MTX 100; steroids 100	NR	NR	NR	NR	NR	NR	NR
Assous 2008[117]	50/86	58 (10)	15 (9)	90	NR	3.5 (1.4)[c]	NR	NR	5.7 (4.2–8.7)[d]	NR	NR	1.9 (0.1–29.2)[d]
Thurlings 2008[118]	24/80	55 (22–75)[d]	12 (1–50)[d]	NR	MTX 100; steroids 100	4.0 (2.0–9.0)[d]	1 (≥1?) (ETN; ADA; IFX)	NR	6.5 (1.1)	NR	37 (22–52)[a]	2.9 (1.2–6.4)[a]
Finckh 2009[136,137]	155/77	55 (13)	12 (9)	88	MTX 67; steroids 58	NR	2 (1–2)[a]	1.60 (1.10–2.00)[a]	5.0 (1.3)	NR	NR	NR
REFLEX extension[139]	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
RTX pooled analysis[139]	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

NR, not reported; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count.
a Median and interquartile range.
b Data presented separately in graphs for 'responders' and 'non-responders'.
c Number of previous DMARDs excluding MTX. All patients had previously been treated with MTX.
d Median (range).
e Baseline data were reported based on all-exposure population and patients receiving ≥two courses RTX, nearly half of whom were TNF naive.
Bold type indicates that the study was an RCT.

- The baseline mean ESR was 48 mm/hour in REFLEX[124–126] and the median value 37 mm/hour and 56 mm/hour in other two studies.
- The mean CRP was 3.7 mg/dl in the REFELEX trial and 3.2 mg/dl in another study; median CRP was 1.9 and 2.9 in the other two studies.

Quality assessment

Randomised controlled trial

The only RCT (REFLEX[124–126]) was of good quality. Full details of the quality assessment are reported in *Table 29*. Randomisation was appropriate and allocation concealment was not described in the paper. Patients and outcome assessors were blinded. It was not clear if data analysts were aware to which group patients were assigned. Withdrawal rate from the RTX group and the placebo group was 18% and 46%, respectively, at week 24, and 63% and 89%, respectively, at week 48. ITT analysis was not used, as 21 patients were excluded from analysis owing to protocol violations.

Non-randomised controlled trials

All the non-RCTs were uncontrolled; four of these were prospective and two were retrospective. Full details of the quality assessment are reported in *Table 30*. All stated clearly their inclusion criteria; however, only in one study was it clear that consecutive patients were included. The percentage of patients withdrawn reported in one study was 25% (at 6 months), the percentage was unclear in two studies and was not applicable in two retrospective studies as only patients with follow-up assessment were included.

REFLEX extension and rituximab pooled analyses

Although some inclusion criteria were stated, in both analyses information on the study characteristics, patient characteristics and methodological appropriateness was insufficient, in particular in the pooled analysis. Details of the quality assessment are reported in *Table 30*.

Results

Tables 31 and *32* present what outcomes were measured in the studies. Outcomes in *Table 31* are reported and described in the main text of this report and those in *Table 32* are reported in *Appendix 10* only. Outcome data from the RTX arm in the RCT are also included in the section on uncontrolled studies for comparison purposes. As data from the REFLEX extension cohort[139]

TABLE 29 Rituximab: RCT quality assessment

Study	Was method of randomisation appropriate?	Was allocation adequately concealed?	Blinding			Patients withdrawn (%)	Was ITT used?	Comments
			Patients	Investigators/ outcome assessors	Data analysts			
REFLEX[124–126]	Yes	Unclear[a]	Yes	Yes[b]	Unclear	Week 24: RTX 18; placebo 46 Week 48:[c] RTX 63; placebo 89	Yes[d]	Twenty-one of the randomised patients were excluded from the ITT population[d]

a Information not described in the papers.
b Blinding of the efficacy assessor was potentially compromised in one of the centres. Patients enrolled in this centre were excluded from ITT analysis.
c Data from the Roche submission.[137]
d A total of 21 patients were excluded from the ITT population, including those for whose treatment was unblinded owing to RTX vial breakage, those who never received treatment, those treated prior to randomisation and those enrolled at a centre where the blinding of the efficacy assessor was potentially compromised. The authors stated that 'sensitivity analyses that included these patients demonstrated no change in the significance of the results'.

Bold type indicates that the study was an RCT.

TABLE 30 Rituximab: non-RCT quality assessment

Study (duration of follow-up)	Study design	Inclusion criteria clearly defined?	Were consecutive patients included in the study?	Patients withdrawn (%)	Comments
Bokarewa 2007[114]	Prospective uncontrolled	Yes	Unclear	NR	
Jois 2007[115]	Prospective uncontrolled	Yes	Unclear	25% at 6 months	
Keystone 2007[116]	Retrospective uncontrolled	Yes	NR	NA	
Assous 2008[117]	Retrospective uncontrolled	Yes	Yes	Unclear	
Thurlings 2008[118]	Prospective uncontrolled	Yes	NR	Unclear	Unclear for those who had subsequent courses at what time point the outcomes were assessed
Finckh 2009[136,137]	Prospective uncontrolled	Yes	NR	NA (only those with follow up assessment were included)	
REFLEX extension[139]	Prospective uncontrolled	Yes	*NA*	NA	
RTX pooled analysis[139]	Retrospective uncontrolled	Unclear	NR	NR	

NA, not applicable; NR not reported.

and the RTX pooled analyses were analysed according to RTX treatment courses, the results of these analyses are described separately from the results of the uncontrolled studies.

Withdrawals

Randomised controlled trial Withdrawal rates are presented in *Figure 29*. At week 24, there were significantly fewer withdrawals for any reason in the RTX arm than in the placebo arm of the REFLEX RCT[124–126] (RR = 0.39, 95% CI 0.29 to 0.51). Risk of withdrawal because of AEs tended to be higher in the RTX than in the placebo group; however, the difference was not statistically significant (RR = 2.71, 95% CI 0.58 to 12.65).

Non-randomised controlled trials Withdrawal rate for any reason at 6 months was reported in only one uncontrolled study[115] and it was 10%. For comparison, 17.9% of patients in the RTX arm in the REFLEX RCT[124–126] withdrew at 6 months for any reason and 2.6% withdrew because of AEs (*Figure 30*). In one study[114] the total number of patients withdrawn by reason was not reported, but it was stated that one patient discontinued RTX treatment after a second infusion (week 4) because of severe headache and stomach pain. Two patients who had a medical history of chronic myocardial ischaemia died of myocardial infarction, one within the first month and the other at 13 months.

ACR20

Randomised controlled trial In the REFLEX trial,[124–126] the percentage of patients who achieved ACR20 response at week 24 in the RTX group was nearly three times that in the placebo group and the difference was statistically significant (RR = 2.85, 95% CI 2.08 to 3.91). At week 48, the response rate based on observed data (of a smaller number of patients) favoured the RTX group, but the difference was not significant (RR = 1.53, 9% CI 0.84 to 2.76); when analysed based on non-responder imputation data, the response rate in the RTX group was nearly five times of that

DOI: 10.3310/hta15140

TABLE 31 Rituximab: outcomes assessed in studies and reported in the main text of the report

Study	Total withdrawal	Withdrawal by reason	ACR (20/50/70)	DAS28	EULAR response	HAQ	QoL	Joint damage	Serious AEs	Infection/ serious infection	Injection/ infusion reaction
REFLEX[124–126]		✓	✓		✓	✓	✓	✓	✓	✓	
Bokarewa 2007[114]									✓	✓	
Jois 2007[115]	✓					✓			✓		
Keystone 2007[116]			✓	Reported graphically	✓	✓					
Assous 2008[117]	✓				✓						
Thurlings 2008[118]				✓	✓				✓	✓	✓
Finckh 2009[136,139]						✓					✓
REFLEX extension[139]	✓	✓	✓	✓	✓	✓					
RTX pooled analysis[139]	✓	✓	✓	✓	✓	✓					

Bold type indicates that the study was an RCT.

TABLE 32 Rituximab: outcomes assessed in studies and reported in the *Appendix 10* only

Study	Other measures of disease activity	Fatigue	Pain	TJC/SJC	CRP/ESR
REFLEX[124–126]	✓	✓	✓	✓	✓
Bokarewa 2007[114]				Reported graphically	Reported graphically
Jois 2007[115]	✓			✓	✓
Keystone 2007[116]					
Assous 2008[117]					✓
Thurlings 2008[118]					
Finckh 2009[136,137]				✓	✓
REFLEX extension[139]					
RTX pooled analysis[139]					

SJC, swollen joint count; TJC, tender joint count.
Bold type indicates that the study was an RCT.

in the placebo group and the difference was significant (RR = 4.92, 95% CI 2.40 to 10.09). Details can be found in *Figure 31*.

Non-randomised controlled trials In the Keystone *et al.*[116] pooled analysis, 24 weeks after the first course of RTX, 65.2% patients had an ACR20 response, while in the RTX arm of the REFLEX trial[124–126] the figure was 51% (*Figure 32*). None of the other uncontrolled studies reported ACR20 responses.

ACR50

Randomised controlled trial At week 24 in the REFLEX trial,[124–126] the percentage of ACR50 responders in the RTX group was nearly five and a half times that of the placebo group and the difference was statistically significant (RR = 5.40, 95% CI 2.87 to 10.16). The effect persisted at week 48, analysed based on either observed data (RR = 4.11, 95% CI 1.06 to 15.85) or non-responder imputation data, and based on non-responder imputation data the response rate in

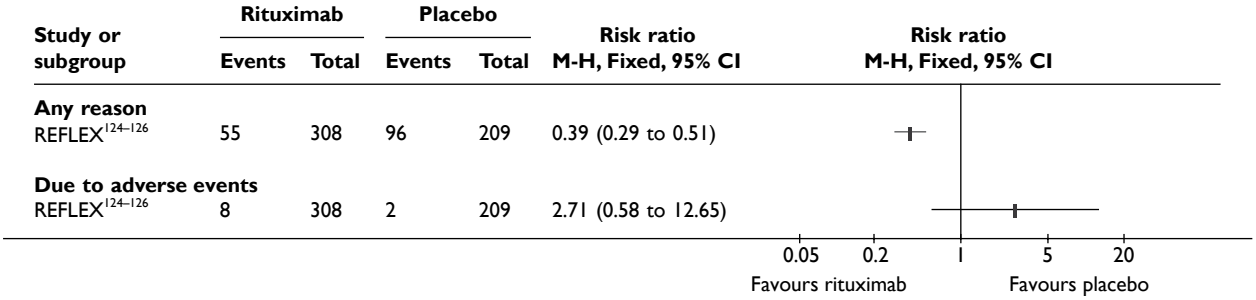


FIGURE 29 Rituximab: withdrawals in the REFLEX RCT[124–126] at 24 weeks by reason.

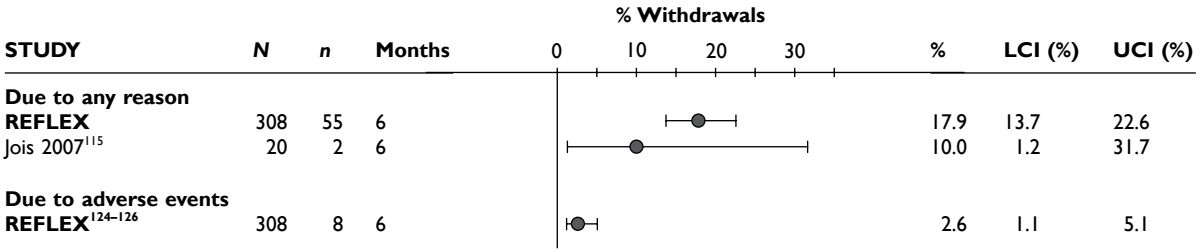


FIGURE 30 Rituximab: withdrawals in uncontrolled studies by reason. LCI, lower confidence interval; UCI, upper confidence interval.

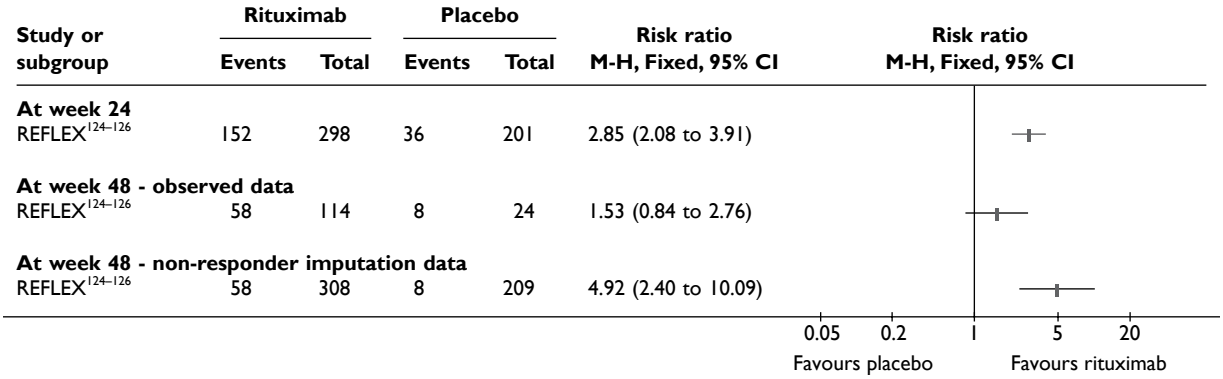


FIGURE 31 Rituximab: ACR20 response in the REFLEX study[124–126] (observed data and non-responder imputation data are from MS).

the RTX group was over 13 times that of the placebo group (RR = 13.23, 95% CI 3.23 to 54.20). Details are presented in *Figure 33*.

Non-randomised controlled trials In the Keystone *et al.*[116] pooled analysis, 24 weeks after the first course of RTX, ACR50 response was observed in 32.9% patients, while in the RTX arm of the REFLEX trial[124–126] it was 26.8% (*Figure 34*). None of the other uncontrolled studies reported ACR50 response.

ACR70

Randomised controlled trial At week 24 the percentage of patients achieving ACR70 response in the RTX group in the REFLEX trial[124–126] was over 12 times of that of the placebo group and the difference was statistically significant (RR = 12.14, 95% CI 2.96 to 49.86). At week 48 the beneficial effect of RTX was not significant based on observed data for a much smaller patient group (RR = 3.37, 95% CI 0.47 to 24.2), but was significant based on non-responder imputation data (RR = 10.86, 95% CI 1.45 to 81.24). See *Figure 35* for details.

Non-randomised controlled trials In the Keystone *et al.* pooled analysis the percentage of ACR70 responders 24 weeks after the first course of RTX was 12.3%; it was similar to that reported in the RTX arm of the REFLEX trial[124–126] (12.1%) (*Figure 36*). No other uncontrolled study reported ACR70 responses.

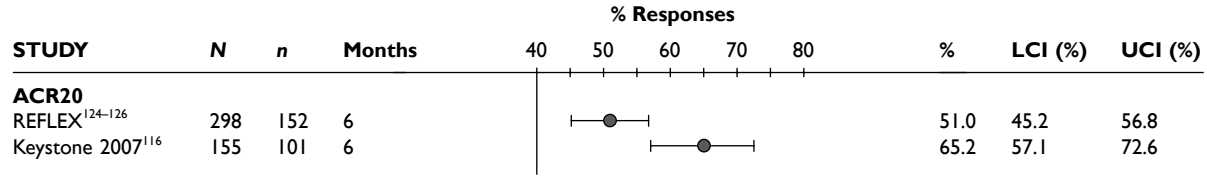


FIGURE 32 Rituximab: ACR20 response in cohorts 24 weeks after first course of RTX. LCI, lower confidence interval; UCI, upper confidence interval.

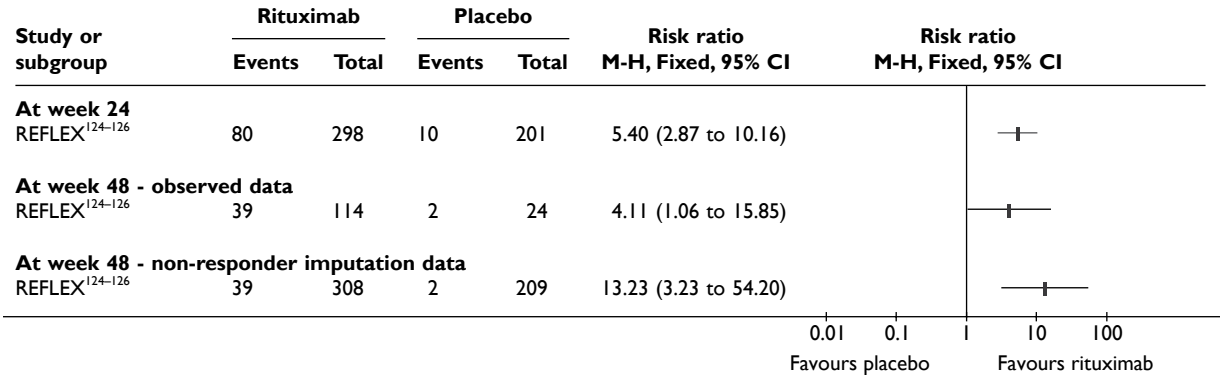


FIGURE 33 Rituximab: ACR50 response in the REFLEX study[124–126] (the observed data and non-responder imputation data were from MS).

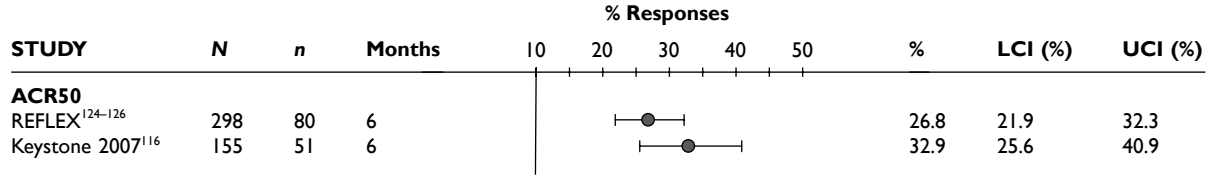


FIGURE 34 Rituximab: ACR50 response in uncontrolled studies 24 weeks after the first course of RTX. LCI, lower confidence interval; UCI, upper confidence interval.

EULAR response

Randomised controlled trial EULAR responses are presented in *Figures 37* and *38*. In the REFLEX trial,[124–126] at week 12 the percentage of patients achieving good or moderate response in the RTX group was over twice that of the placebo group, as was the percentage achieving a good response; the effects were statistically significant (RR = 2.02, 95% CI 1.64 to 2.49 and RR = 2.23, 95% CI 1.12 to 4.41, respectively). At week 24 the percentage of patients achieving a EULAR good or moderate response in the RTX group was nearly three times that of the placebo group and the effect was significant (RR = 2.96, 95% CI 2.25 to 3.89); the rate of achieving a good response was also higher in the RTX group and the difference was statistically significant (RR = 7.59, 95% CI 2.77 to 20.77).

Non-randomised controlled trials None of the uncontrolled studies reported EULAR response at 3 months, whereas three reported it at 6 months. At 3 months in the RTX arm of the REFLEX RCT,[124–126] 68.5% of patients had moderate or good response, with 11.1% having achieved a good response; at 6 months the rates remained similar (64.8% and 15.1%, respectively). At 6 months the percentage of good or moderate EULAR responders in four uncontrolled studies including

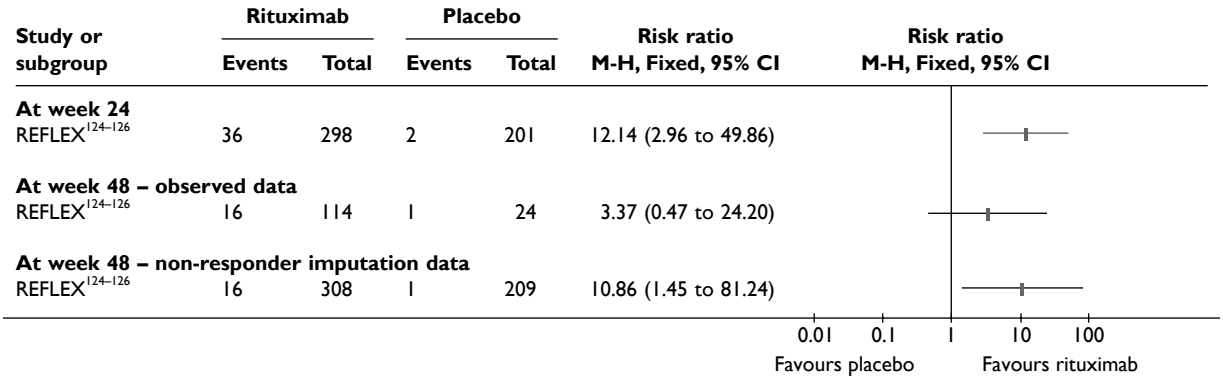


FIGURE 35 Rituximab: ACR70 response in the REFLEX study[124–126] (those based on observed data and non-responder imputation data were from manufacturer's submission).

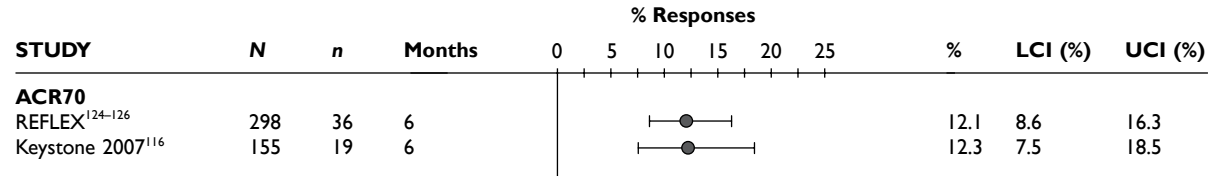


FIGURE 36 Rituximab: ACR70 response in uncontrolled studies 24 weeks after the first course of RTX.

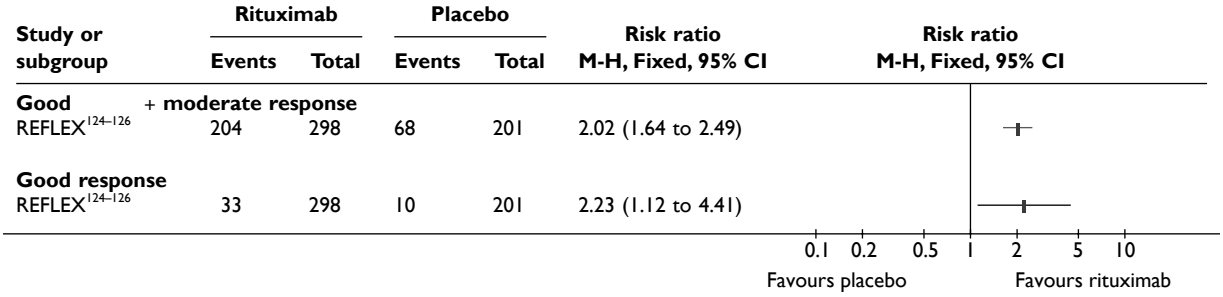


FIGURE 37 Rituximab: EULAR response at week 12 in the REFLEX study[124–126] (data from the manufacturer's submission).

the RTX arm of the REFLEX trial[124–126] ranged from 64.8% to 82%, and the good response rate ranged from 15.1% to 36%. The REFLEX trial[124–126] had the lowest percentage of responders in both categories. One study also reported EULAR low disease activity and remission at 6 months (13.3% and 5.7%, respectively). See *Figure 39* for details.

DAS28

Randomised controlled trial In the REFLEX trial,[124–126] at week 24, the RTX arm had a significantly smaller mean DAS28 score and significantly greater reduction in the mean DAS28 score from baseline than the placebo arm (–1.40, 95% CI –1.67 to –1.13, and – 1.50, 95% CI –1.74 to –1.26, respectively). At week 24, the proportion of patients with DAS28 improvement in the RTX group was over five times that in the placebo group and the difference was statistically significant. See *Figures 40–42* for details.

Non-randomised controlled trials DAS28 score at 3 months was available in only one uncontrolled study (median DAS28 = 5.60). DAS28 score at 6 months was measured in three studies. The mean score was 5.0 in one study and it was the same as that of the RTX arm of the REFLEX trial.[124–126] Two studies provided a median score and it was 5.50 and 3.97. See *Figure 43* for details. (Note: for the Jois *et al.*[115] and Assous *et al.*[117] studies scores were reported as medians.)

FIGURE 38 Rituximab: EULAR response at week 24 in the REFLEX study.[124–126]

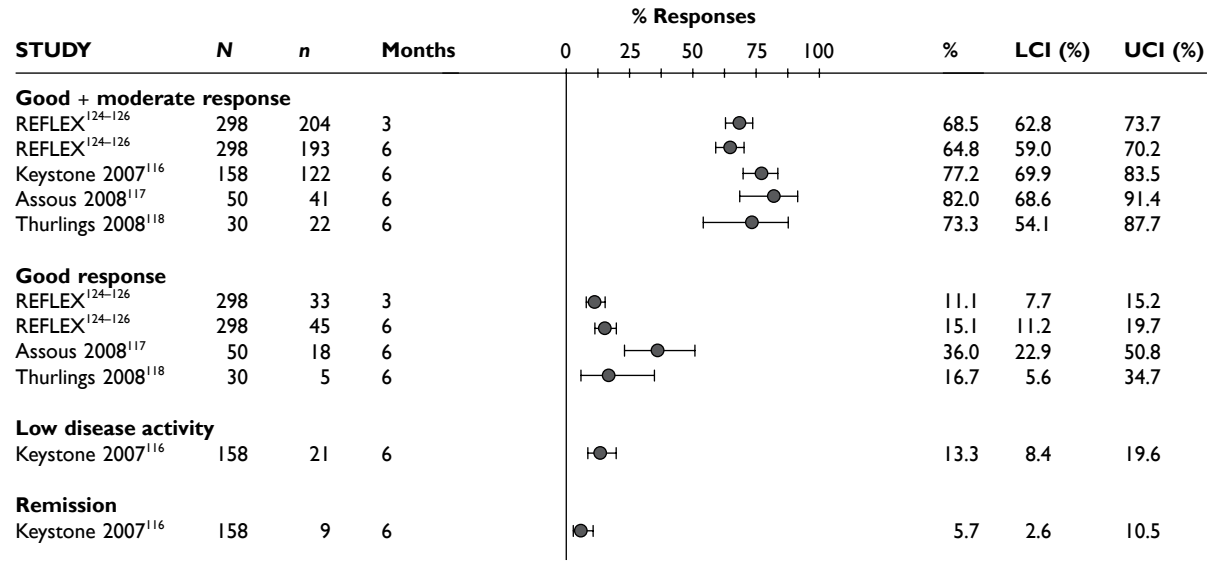


FIGURE 39 Rituximab: EULAR response in uncontrolled studies (3-month data for the REFLEX trial[124–126] from MS). LCI, lower confidence interval; UCI, upper confidence interval.

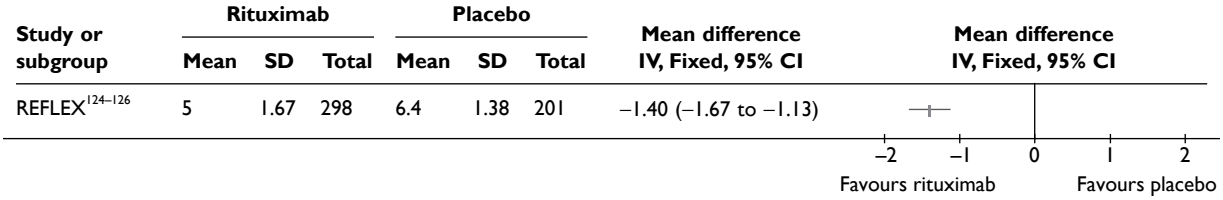


FIGURE 40 Rituximab: DAS28 score at week 24 in the REFLEX trial[124–126] (last observation carried forward, data from MS). SD, standard deviation.

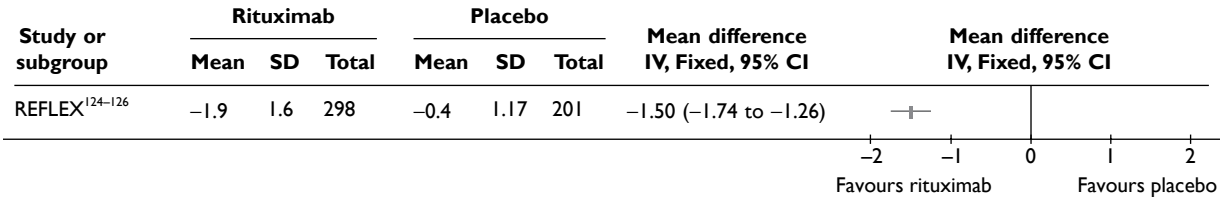


FIGURE 41 Rituximab: DAS28 score change from baseline at week 24 in the REFLEX trial[124–126] (last observation carried forward, data from MS). SD, standard deviation.

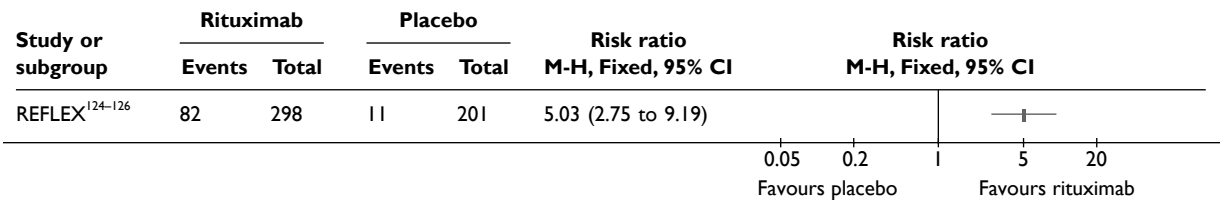


FIGURE 42 Rituximab: percentage of patients with DAS28 improvement from baseline at week 24 in the REFLEX trial[124–126] (last observation carried forward, data from the MS).

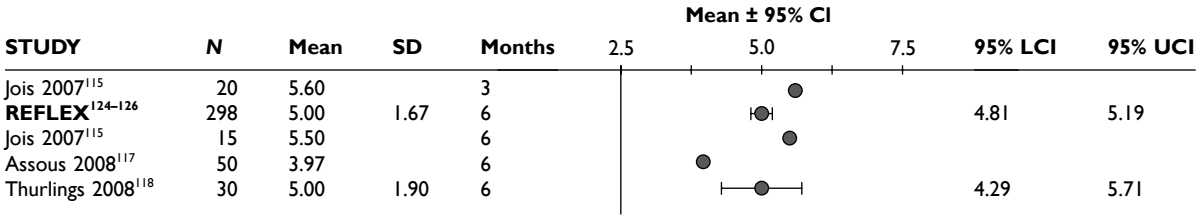


FIGURE 43 Rituximab: mean DAS28 in uncontrolled studies. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

In Finckh *et al.*[136] the change in mean DAS28 score from baseline at 6 months was reported only for a subgroup of 50 patients. It was similar to that reported for the RTX arm of the REFLEX trial[124–126] and both showed significant improvement (–1.90, 95% CI –2.08 to –1.72, and –1.61, 95% CI –1.98 to –1.24, respectively). See *Figure 44* for details.

Health Assessment Questionnaire

Randomised controlled trial In the REFLEX trial,[124–126] the RTX group had significantly more reduction in mean HAQ score from baseline at week 24 compared with the placebo group (mean difference = –0.30, 95% CI –0.40 to –0.20; *Figure 45*).

The percentage of patients who showed HAQ improvement, defined as a decrease in score from baseline of greater than 0.25, in the RTX group of the REFLEX trial[124–126] was nearly twice that

of the placebo group at week 12, and over two and a half times as high at week 24; both effects were statistically significant (RR = 1.63, 95% CI 1.29 to 2.07, and RR = 2.55, 95% CI 1.89 to 3.43, respectively). See *Figure 46* for details.

At week 24, the observed percentage of patients with minimal clinically meaningful improvement in HAQ, defined as a decrease in HAQ score of 0.22, in the RTX group of the REFLEX trial,[124–126] was over 1.6 times that of the placebo groups and the difference was significant; whereas observed at week 48 there was no significant difference (*Figure 47*).

When analysed based on non-responder imputation data, the percentage of patients with minimal clinically meaningful improvement in HAQ at week 24 and week 48 in the RTX group was over two and a half and over three and a half times that of the placebo group (58% vs 23% and 23% vs 6%, respectively) and both differences were statistically significant (*Figure 48*).

Non-randomised controlled trials Two uncontrolled studies reported HAQ score. The median HAQ score in one study[115] was 2.13 (range 0.63–2.88) at 3 months and decreased to 1.86 (range 1–3) at 6 months; however, in both cases, the reduction compared with baseline was not significant. In the Keystone *et al.* study,[116] the percentage of patients with a decrease in the mean HAQ score of greater than or equal to 0.22 from baseline at week 24 (after one course of RTX treatment) was 71.8%, which is very similar to the observed rate reported in the RTX arm of the REFLEX trial[124–126] (70.5%) (*Figure 49*).

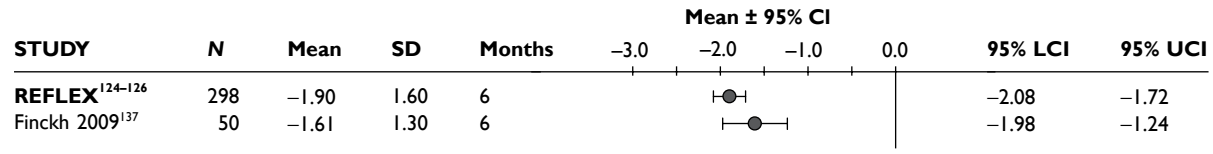


FIGURE 44 Rituximab: DAS28 scores change from baseline in uncontrolled studies (data for REFLEX[124–126] from MS). LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

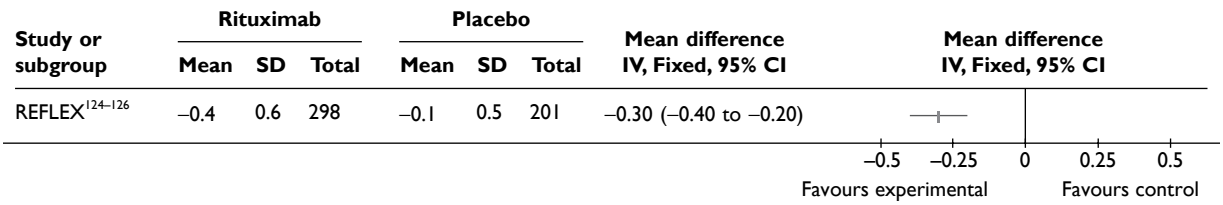


FIGURE 45 Rituximab: mean change in HAQ scores from baseline at week 24 in REFLEX trial.[124–126] SD, standard deviation.

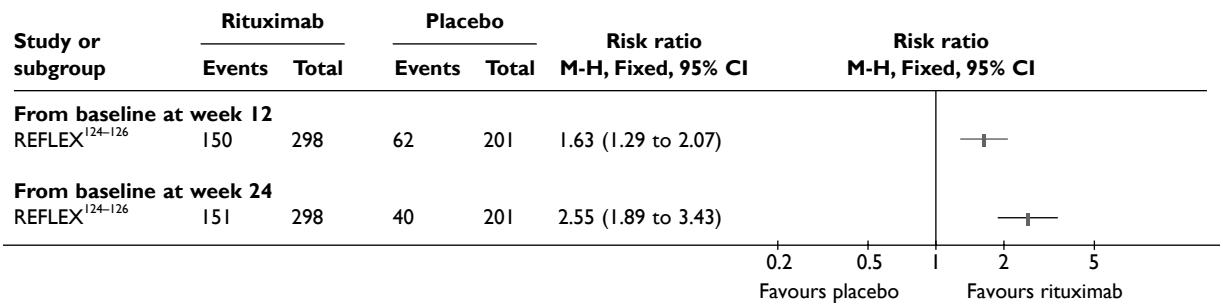


FIGURE 46 Rituximab: percentage of patients with a decrease in HAQ score > 0.25 from baseline in the REFLEX study[124–126] (data from MS).

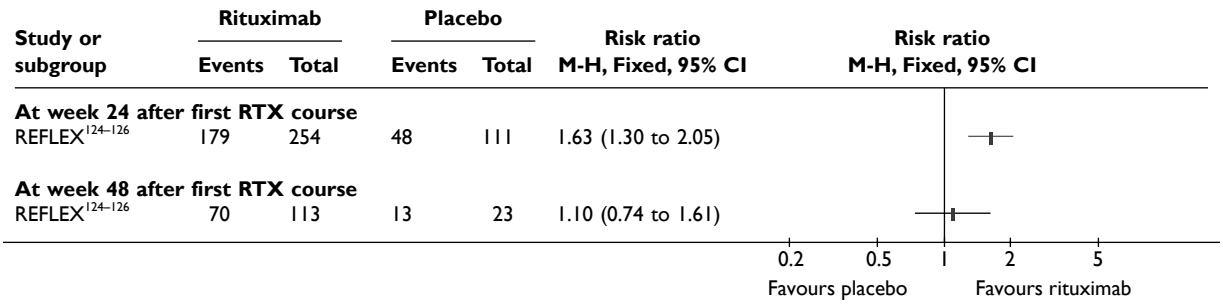


FIGURE 47 Rituximab: percentage of patients with a clinically meaningful improvement in HAQ, 24 and 48 weeks after the first course of RTX (observed data from MS).

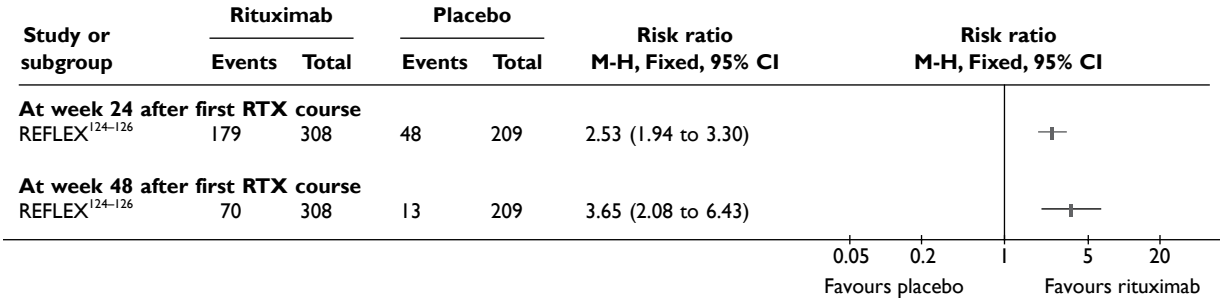


FIGURE 48 Rituximab: percentage of patients with a clinically meaningful improvement in HAQ, 24 and 48 weeks after the first course of RTX (non-responder imputation data from MS).

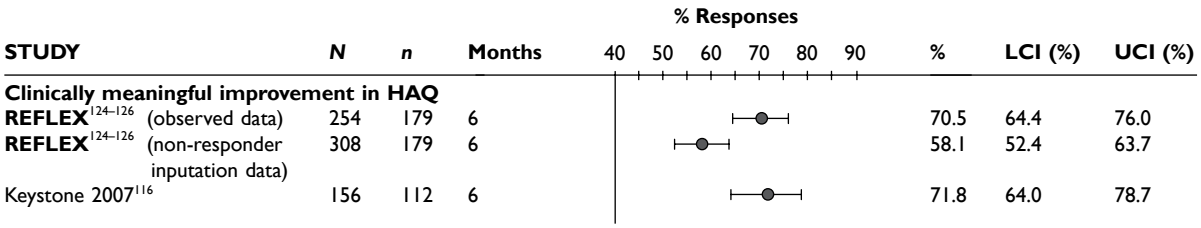


FIGURE 49 Rituximab: percentage of patients with clinically meaningful improvement in HAQ score from baseline at week 24. LCI, lower confidence interval; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

Joint damage

Randomised controlled trial The RTX group of the REFLEX trial[124–126] had significantly less changes in Sharp-Genant total score from baseline than the placebo group at both week 56 (mean difference = –1.12, 95% CI –2.13 to –0.11) and week 104 (mean difference = –1.67, 95% CI –2.67 to –0.67). At week 56 the percentage of patients with no worsening of Sharp-Genant total score from baseline in the RTX group was nearly one and a half times that in the placebo group and the difference was statistically significant. Sharp-Genant total score measured at week 104 favoured the RTX group but the difference was not statistically significant (mean difference = –3.53, 95% CI –9.21 to 2.15). See *Appendix 10* for details.

There was significantly less change from baseline in the erosion score in the RTX group than in the placebo group at week 56 (mean difference = –0.75, 95% CI –1.43 to –0.07), and at week 104 the significant difference became larger (mean difference = –1.08, 95% CI –1.73 to –0.43). The erosion score at week 104 favoured the RTX arm, but the difference was not statistically significant (mean difference = –2.48, 95% CI–5.55 to 0.59). The percentage of patients with no erosive progression from baseline at week 104 in the RTX group was nearly one and a half times

that of the placebo group and the difference was statistically significant (RR = 1.38, 95% CI 1.14 to 1.66).

Joint space narrowing score change from baseline was smaller in the RTX group than in the placebo group both at week 56 and week at 104; the difference was not statistically significant at week 56 but became significant at week 104, though at week 104 the joint space narrowing score was not significantly lower in the RTX group than in the placebo group.

Non-randomised controlled trials None of the uncontrolled studies reported joint damage.

Quality of life

Randomised controlled trial Mean Short Form questionnaire-36 items (SF-36) mental and physical health scores measured at week 24 in the REFLEX trial[124–126] were both significantly higher in the RTX group than in the placebo group (*Figure 50*). The RTX group increased mean SF-36 physical health score by 5.16 and mean SF-36 mental health score by 3.07 higher than in the placebo group, and the differences were statistically significant (*Figure 51*).

Non-randomised controlled trials None of the uncontrolled studies reported QoL.

Serious adverse events

Randomised controlled trial In the REFLEX trial,[124–126] the percentage of patients with serious AEs was lower in the RTX group than in the placebo group; the difference was not statistically significant (RR = 0.74, 95% CI 0.42 to 1.31). See *Figure 52* for details.

Non-randomised controlled trials In one 12-month study,[114] one patient (2%) had severe headache and stomach pain 1 day after RTX infusion and this led to a discontinuation of treatment. A 6-month study,[115] stated that no major side effects were found during the study. During a 6-month period the Thurlings *et al.*[118] study reported five serious AEs (16.7%): two severe

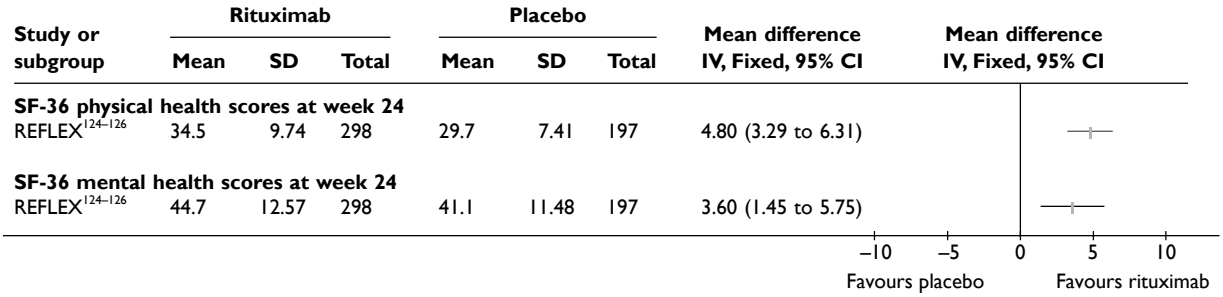


FIGURE 50 Rituximab: mean SF-36 items scores at week 24 in the REFLEX trial[124–126] (last observation carried forward, data from MS). SD, standard deviation.

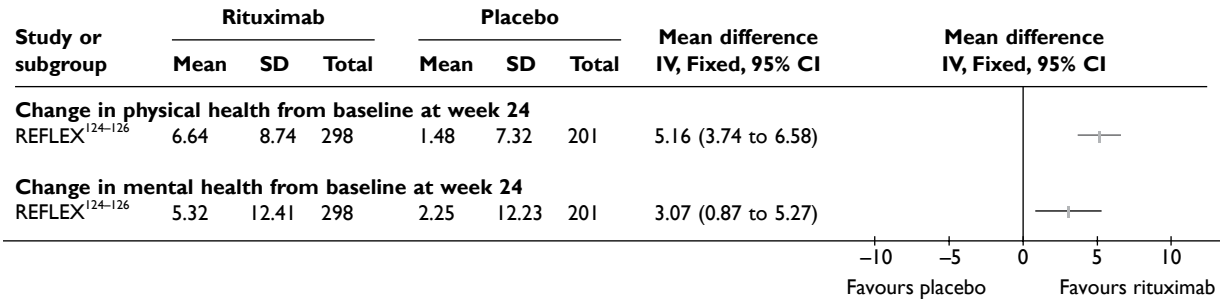


FIGURE 51 Rituximab: change in SF-36 items scores from baseline to week 24 in REFLEX trial.[124–126] SD, standard deviation.

infusion reactions, one arterial embolism, one pulmonary embolism and one toxic hepatitis. The other studies did not report information on serious AEs.

Any infection/serious infection

Randomised controlled trial In the REFLEX trial[124–126] both the percentage of patients with any infections and the percentage of patients with serious infections were greater in the RTX group than in the placebo group; however, none of the differences was statistically significant (RR = 1.08, 95% CI 0.87 to 1.35 and RR = 1.58, 95% CI 0.41 to 6.05, respectively). See *Figure 53* for details.

Non-randomised controlled trials In the Bokarewa *et al.* study[114] 3 months after the treatment with RTX, pneumonia requiring hospitalisation was reported in one patient (2.0%). In Thurlings *et al.*[118] the incidence of infection per patient-year was 0.9: 48 infections requiring antibiotic, antimycotic, or antiviral treatment and one serious infection requiring i.v. antibiotics occurred among 30 patients over 2 years of follow-up. One serious infection requiring i.v. antibiotics was observed in this study.

Injection site reaction/infusion reaction

Randomised controlled trial In the REFLEX trial,[124–126] the percentage of patients with acute infusion reactions did not differ significantly between groups (RR = 1.24, 95% CI 0.90 to 1.83, for the first course and RR = 0.74, 95% CI 0.43 to 1.24, for the second course). See *Figure 54* for details.

Non-randomised controlled trials One study (Finckh *et al.*,[137] subgroup of 50 patients) reported three mild-to-moderate infusion reactions. Another study[118] reported two severe infusion reactions. The other studies did not report information on infusion site reactions.

Data reported by treatment course

Pooled analysis (data from Keystone et al.*)*

In the Keystone *et al.* study,[116] based on evaluable data, the percentage of patients achieving ACR responses increased from course 1 to course 2 of RTX measured 24 weeks after each course (*Figure 55*). A similar pattern was seen for the percentage of patients with EULAR response 24 weeks after course 1 and course 2 (*Figure 56*).

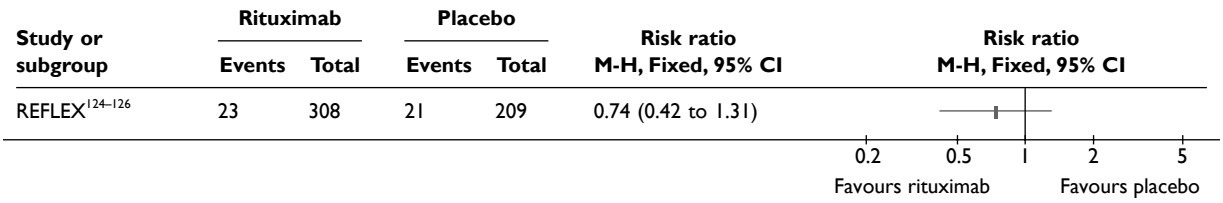


FIGURE 52 Rituximab: serious AEs at week 24 in the REFLEX trial.[124–126]

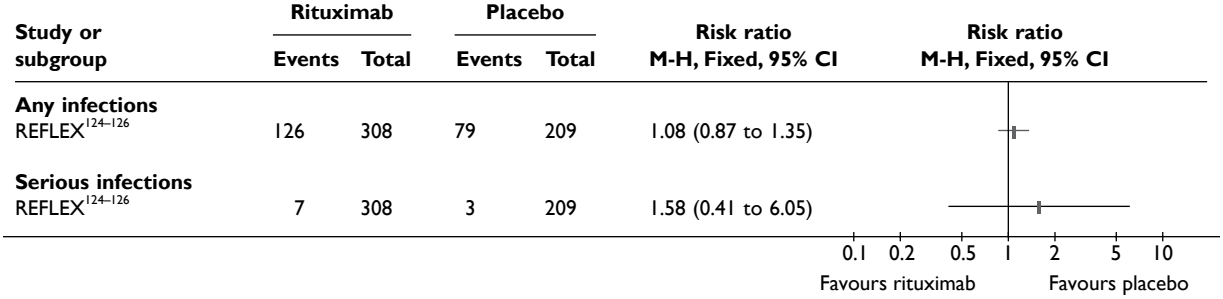


FIGURE 53 Rituximab: any infection and serious infection at week 24 in the REFLEX trial.[124–126]

The percentage of patients who achieved meaningful improvement in HAQ, i.e. had a decrease of HAQ scores at least 0.22 from baseline, were similar 24 weeks after course 1 and course 2 of RTX treatment (*Figure 57*).

Data from manufacturer's submission

Data analysis based on the MS can be found together with all additional analyses in *Appendix 10*.

Summary

For the assessment of effectiveness of RTX in comparison with standard care, one RCT and six uncontrolled studies were identified. Follow-up duration ranged from 3 months to 24 months. Patients included in the studies were generally similar. The main results of the seven studies are summarised in *Table 33*.

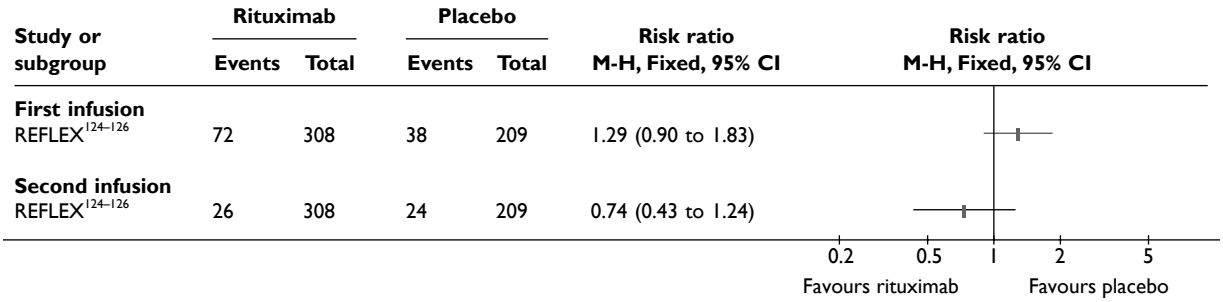


FIGURE 54 Rituximab: percentage of patients who had acute infusion reactions after the first and second infusion.

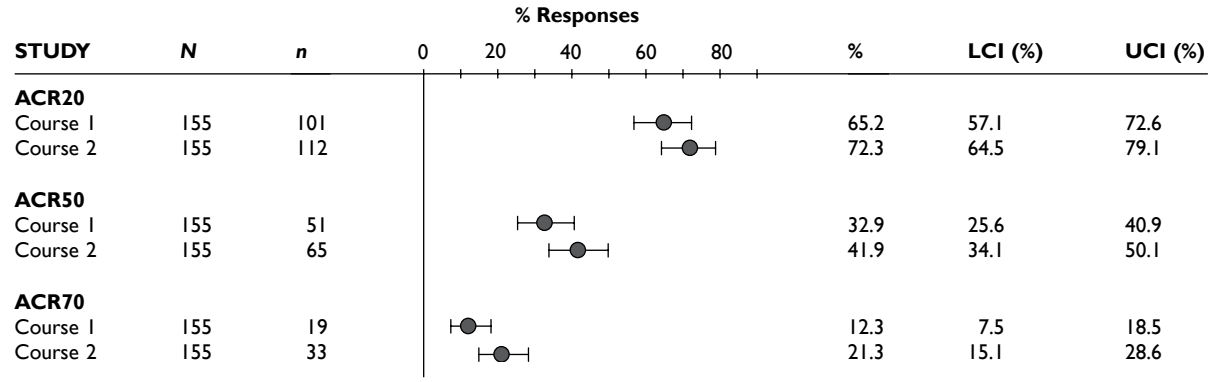


FIGURE 55 Percentage of patients achieving ACR responses 24 weeks after course 1 and course 2 – based on evaluable patients who had prior TNF inhibitor. LCI, lower confidence interval; UCI, upper confidence interval.

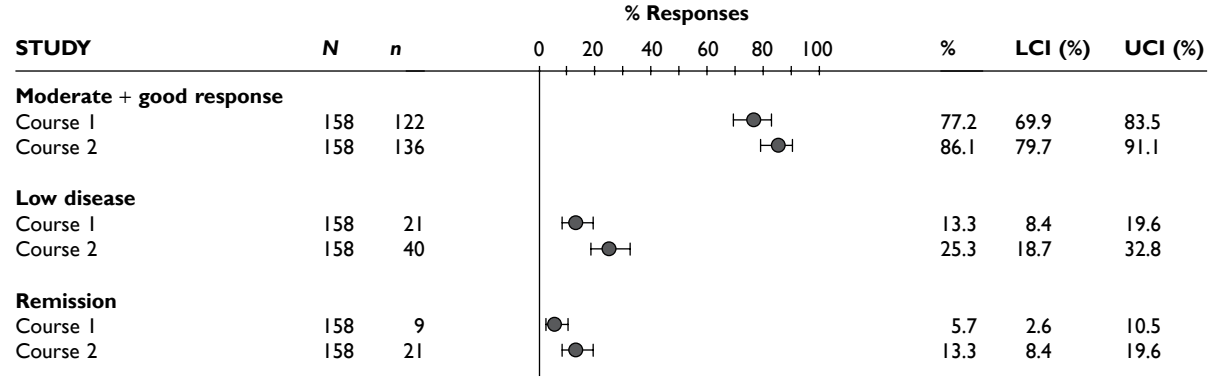


FIGURE 56 Percentage of patients with EULAR responses 24 weeks after course 1 and course 2 – based on evaluable patients who had prior TNF inhibitor. LCI, lower confidence interval; UCI, upper confidence interval.

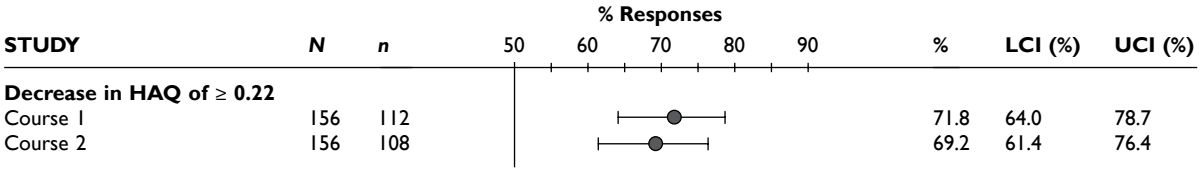


FIGURE 57 Percentage of patients with a decrease in HAQ score of ≥0.22 at week 24 after course 1 and course 2 – based on evaluable patients who had prior TNF inhibitor. LCI, lower confidence interval; UCI, upper confidence interval.

TABLE 33 Rituximab: summary of main results

	RCT [result (95% CI)]		Uncontrolled studies	
Outcome	**6 months (RTX vs placebo)**	**6 months (RTX arm)**	**3 months**	**6 months**
Withdrawals (%):				
■ for any reason	RR = 0.39 (0.29 to 0.51), favours RTX	17.9	NR	10.0
■ due to lack of efficacy	NR	NR	NR	NR
■ due to AEs	RR = 2.71 (0.58 to 12.65), NS	2.6	NR	NR
ACR20 response (%)	RR = 2.85 (2.08 to 3.91), favours RTX	51.0	NR	65.2
ACR50 response (%)	RR = 5.40 (2.87 to 10.16), favours RTX	26.8	NR	32.9
ACR70 response (%)	RR = 12.14 (2.96 to 49.86), favours RTX	12.1	NR	12.3
EULAR response (%):				
■ good and moderate response	RR = 2.96 (2.25 to 3.89), favours RTX	64.8	NR	73.3–82.0
■ good response	RR = 0.76 (0.52 to 1.12), NS	15.1	NR	16.7–36.0
DAS28: mean change from baseline	Mean difference = –1.40 (–1.67 to –1.13), favours RTX	–1.90	NR	–1.61
HAQ: mean change from baseline	Mean difference = –0.30 (–0.40 to –0.20), favours RTX	–0.40	NR	NR
Patients with an improvement in HAQ > 0.25 from baseline (%)	RR = 2.55 (1.89 to 3.43), favours RTX	50.7	NR	71.8
Joint damage (Sharp–Genant total score)	Mean difference (week 56) = –1.12 (–2.13 to –0.11), favours RTX	0.66 (week 56)	NR	NR
QoL				
Change from baseline in SF-36 physical health score	Mean difference = 5.16 (3.74 to 6.58), favours RTX	6.64	NR	NR
Change from baseline in SF-36 mental health score	Mean difference = 3.07 (0.87 to 5.27)	5.32	NR	NR
Serious AEs (%)	RR = 0.74 (0.42 to 1.31), NS	7.5	NR	0–16.7 (2% for 12 months)
Any infections (%)	RR = 1.08 (0.87 to 1.35), NS	40.9	NR	Infections (requiring antibiotic, antimycotic or antiviral treatment) per patient-year = 0.9 (over 2 years)
Serious infections (%)	RR = 1.58 (0.41 to 6.05), NS	2.3	2	NR
Infusion reaction (%)				
First infusion reaction	RR = 1.29 (0.90 to 1.83), NS	23.4	NR	NR
Second infusion reaction	RR = 0.74 (0.43 to 1.24), NS	8.4	NR	NR

NR, not reported; NS, not significant.

Abatacept

Overview of evidence

Three studies were identified that assessed ABT: one RCT [abatacept trial in treatment of anti-TNF inadequate responders (ATTAIN[127–132])], an extension of this RCT (ATTAIN LTE[119]) and an uncontrolled study [abatacept researched in rheumatoid arthritis patients with an inadequate anti-TNF response to validate effectiveness (ARRIVE)[120]].

Patients were included in the ATTAIN LTE[119] after completing 6 months of the RCT. It was reported that in total 74.4% of the placebo group and 86.4% of the ABT group were included in the extension.

Patients in the studies were non-responders to at least one TNF inhibitor. In the ATTAIN RCT[127–132] and LTE[119] lack of efficacy was the primary reason for switching biologic agents. In ARRIVE[120] patients discontinued the previous TNF inhibitor because of lack of efficacy, safety concerns or intolerability.

All studies were carried out in North America and Europe. ARRIVE[120] additionally included Mexican patients. No information was provided if these studies included UK patients. Follow-up was 6 months for the ATTAIN RCT[127–132] and ARRIVE study.[120] In the ATTAIN LTE[119] patients were followed up for up to 5 years; however, there was no published data beyond 2 years. Further details are provided in *Table 34*.

Patient characteristics

Full details of patient characteristics are reported in *Table 35*.

TABLE 34 Abatacept: characteristics of included studies

Study	Country	Design	Reason for switching	Prior TNF inhibitors; *n*	Treatment arms (no. of patients)	Duration of follow-up	Comments
RCTs							
ATTAIN[127–132]	North America and Europe	Parallel prospective	Primarily lack of efficacy	Any; 1–2	ABT (258) PL (133)	6 months	
Non-randomised comparative studies							
None were identified							
Uncontrolled studies							
ATTAIN LTE[119]	North America and Europe	Uncontrolled prospective LTE of RCT	Primarily lack of efficacy	Any; 1–2	ABT (317)	Up to 5 years	Some patients have not yet completed the 5-year follow-up; published data only up to 2 years; data beyond that from MS
ARRIVE[120]	USA, EU, Mexico	Uncontrolled prospective	Lack of efficacy, safety, intolerability	Any; 1–3	ABT (1,046)	6 months	Two main subgroups: patients switched to ABT after a washout period and those who switched directly

PL, placebo.
Bold type indicates that the study was an RCT.

TABLE 35 Abatacept: baseline patient characteristics

Study	Number of patients/% female	Age (years), mean (SD)	RA duration (years), mean	RF positive (%)	% of patients on concomitant DMARDs and steroids	Number of previous DMARDs, mean (SD)	Number of previous TNF inhibitors, mean (SD)	HAQ, mean (SD)	DAS28, mean (SD)	TJC/SJC, mean (SD)	ESR (mm/hour), mean (SD)	CRP (mg/dl), mean (SD)
ATTAIN[127–132]	391/78	53.2 (12.0)	11.9 (8.6)	73.2	MTX (77.8); HCQ (8.9); LEF (8.7); sulfasalazine (8.0); corticosteroids (68.3)	NR	1–2	1.8 (0.6)	6.5 (0.9)	31.7 (13.1) of 68/22.2 (10.1) of 66	NR	4.4 (3.9)
ATTAIN LTE[119]	317/78	53.0 (11.7)	11.8 (8.6)	NR	Continued MTX, DMARDs and corticosteroids allowed	NR	1–2	1.8 (0.6)	6.5 (0.8)	31.8 (13.4)/22.3 (10.4)	NR	4.2 (3.7)
ARRIVE[120]	1,046/81	54.4 (12.4)	11.6 (9.5)	61.3	MTX (69.8), AZA (4.1), GST (0.5), HCQ/chloroquine (15.0), LEF (12.8), sulfasalazine (8.8), corticosteroids (58.4)	NR	1–3	1.7 (0.6)	6.2 (0.7)	17.8 (6.0)/13.6 (5.5)	NR	2.1 (3.0)

NR, not reported; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count.
Bold type indicates that the study was an RCT.

TABLE 36 Abatacept: RCT quality assessment

			Blinding					
Study	Was randomisation appropriate?	Was allocation adequately concealed?	Patients	Investigators/outcome assessors	Data analysts	Patients withdrawn (%)	Was ITT used?	Comments
ATTAIN[127–132]	Yes	Yes	Yes	Yes	Unclear	ABT 13.6; PL 25.6	No; modified ITT used (patients who were given at least one dose of the drug)	Two patients excluded from analysis because of protocol violation

Bold type indicates that the study was an RCT.

The number of patients included in the studies was 391 in the ATTAIN RCT,[127–132] 317 in its LTE[119] and 1,046 in the ARRIVE study.[120] Patient characteristics were generally similar across studies and study arms:

- The percentage of female patients ranged from 78% to 81%.
- The mean age ranged from 53.0 to 54.4 years.
- The mean disease duration ranged from 11.6 to 11.9 years.
- In two studies the percentage of RF-positive patients ranged from 61.3% to 73.2%; it was not reported in the ATTAIN LTE.[119]
- Concomitant DMARDs were reported in detail in ATTAIN[127–132] and ARRIVE:[120] 69.8%–77.7% patients were on MTX; other DMARDs included HCQ (8.9%–15.0%), LEF (8.7%–2.8%) and sulfasalazine (8.0%–8.8%). In the ARRIVE study,[120] AZA (4.1%) and gold (0.5%) were also used.
- In two studies 58.4%–68.3% of patients were receiving corticosteroids; this information was not reported in detail in the ATTAIN LTE.[120]
- The number of previously used conventional DMARDs was not reported in any of the studies.
- The number of previous TNF inhibitors ranged from one to two in the ATTAIN[127–132] and ATTAIN LTE[119] studies and from one to three in the ARRIVE study.[120]
- The mean baseline HAQ ranged from 1.7 to 1.8.
- The mean DAS28 score ranged from 6.2 to 6.5.
- The mean number of tender and swollen joints ranged from 17.8 to 31.8 and from 13.6 to 22.3, respectively.
- Baseline ESR was not reported in any of the studies.
- CRP ranged from 2.1 mg/dl to 4.4 mg/dl.

Quality assessment

Randomised controlled trial The only RCT (ATTAIN[127–132]) was of high quality. Full details of the quality assessment are reported in *Table 36*. Randomisation and allocation concealment were appropriate. Patients and investigators/outcome assessors were blinded. It was not clear if data analysts knew to which group patients were assigned. A total of 13.6% of patients were withdrawn from the ABT group and 25.6% from the placebo group. ITT analysis was not used, as only data from patients who received at least one dose of the study drug were analysed. Two patients were excluded from analysis because of protocol violations, possibly post hoc. The potential impact on the results is likely to be small.

Non-randomised controlled trials Both non-randomised studies were uncontrolled and prospective. Full details of the quality assessment are reported in *Table 37*. Both studies stated clearly their inclusion criteria; however, it was not clear if consecutive patients were included in ARRIVE.[120] The percentage of patients withdrawn from the study was 18% in the ARRIVE study[120] at 6 months and 30% in the ATTAIN LTE[119] at 2 years.

TABLE 37 Abatacept: non-RCT quality assessment

Study	Study design	Inclusion criteria clearly defined?	Were consecutive patients included in the study?	Patients withdrawn (%)	Comments
ATTAIN LTE[119]	Uncontrolled long-term open-label extension of RCT	Yes	NA	30	Data for 2-year follow-up
ARRIVE[120]	Uncontrolled prospective	Yes	Unclear	18	

NA, not applicable.

Results

The RCT and non-randomised studies were analysed separately. Data from the ABT arm of the ATTAIN RCT are included in all figures referring to uncontrolled studies for comparison.

Table 38 indicates which of the outcomes reported in the main text of the report were assessed in individual studies and *Table 39* provides similar information for outcomes described in *Appendix 10* only.

Withdrawals

Randomised controlled trial There were significantly fewer withdrawals for any reason in the ABT arm than in the placebo arm of the ATTAIN RCT[127–132] (RR = 0.53, 95% CI 0.35 to 0.81). There were also significantly fewer withdrawals in the ABT group because of lack of efficacy (RR = 0.27, 95% CI 0.15 to 0.49). The risk of withdrawal because of AEs was similar in both groups (RR = 0.93, 95% CI 0.32 to 2.71). Details of the analysis are presented in *Figure 58*.

Non-randomised control trials At 6 months 17.8% patients withdrew from the ARRIVE study.[120] This percentage was slightly higher than in the ABT-treated arm of the RCT. At 2 years, 30% of patients had withdrawn from the ATTAIN LTE.[119] In both studies, more patients withdrew because of lack of efficacy than because of AEs. A similar relationship was observed in the ABT arm of the RCT. Full details are presented in *Figure 59*.

ACR20 response

Randomised control trial ATTAIN[127–132] reported ACR20 response at 3 and 6 months. At both follow-up times the risk of an ACR20 response was over two and a half times higher in the ABT group than in the placebo group and the difference was statistically significant (for 3 months, RR = 2.53, 95% CI 1.72 to 3.73; for 6 months, RR = 2.56, 95% CI 1.77 to 3.69). Details can be found in *Figure 60*.

Non-randomised control trials Of the uncontrolled studies, only the ATTAIN LTE[119] reported ACR20 response. Results are reported by subgroup based on whether patients were originally randomised to ABT or placebo in the randomised phase (see *Figure 61*). After 6 months of ABT treatment, 57.3% patients in the group initially randomised to ABT and 63.6% in the group initially randomised to placebo achieved an ACR20 response. This was slightly more than in the ABT arm of the RCT (50.0%). After 6 months, there was a further increase in the percentage of ACR20 responders at 12 months in those initially randomised to ABT followed by a decrease up to 5 years (30.3%). Among those initially randomised to placebo, there was a decrease in the percentage of responders from 12 months onwards, and at 54 months 30.3% of patients were ACR20 responders.

If only patients for whom data were available at different time points were analysed, the increase in percentage of ACR20 responders continued to 3 years (82.1%) and then decreased to 65.6% at 5 years for patients initially randomised to ABT. In the same analysis, among patients initially randomised to placebo there was an increase in the percentage of ACR20 responders up to 42 months (82.0%), and at 54 months 78.9% were ACR20 responders.

ACR50 response

Randomised controlled trial At 6 months the percentage of ACR50 responders was over five times higher in the ABT group than in the placebo group of the ATTAIN trial[127–132] and the difference was statistically significant (RR = 5.36, 95% CI 2.19 to 13.10). Details are presented in *Figure 62*.

Non-randomised contolled trials Of the uncontrolled studies, only the ATTAIN LTE[119] reported ACR50 response. Results are reported by subgroup based on whether patients were originally

DOI: 10.3310/hta15140

TABLE 38 Abatacept: outcomes assessed in studies and reported in the main text of the report

Study	Total withdrawal	Withdrawal by reason	ACR (20/50/70)	DAS28	EULAR response	HAQ	QoL	Joint damage	Serious AEs	Infection/ serious infection	Injection/ infusion reaction
ATTAIN[127–132]	✓	✓	✓	✓		✓	✓		✓	✓	✓
ATTAIN LTE[119]	✓	✓	✓	✓		✓			✓	✓	
ARRIVE[120]	✓	✓		✓			✓		✓	✓	✓

Bold type indicates that the study was an RCT.

TABLE 39 Abatacept: outcomes assessed in studies and reported in *Appendix 10* only

Study	Other measures of disease activity	Fatigue	Pain	TJC/SJC	CRP/ESR
ATTAIN[127–132]			✓		
ATTAIN LTE[119]	✓	✓	✓		
ARRIVE[120]					

SJC, swollen joint count; TJC, tender joint count.
Bold type indicates that the study was an RCT.

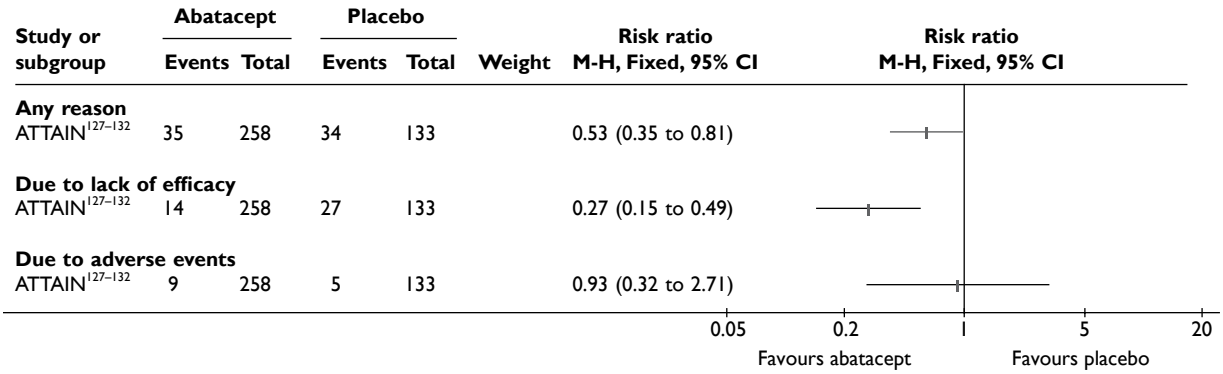


FIGURE 58 Abatacept: withdrawals by reason in the ATTAIN RCT[127–132] at 6 months.

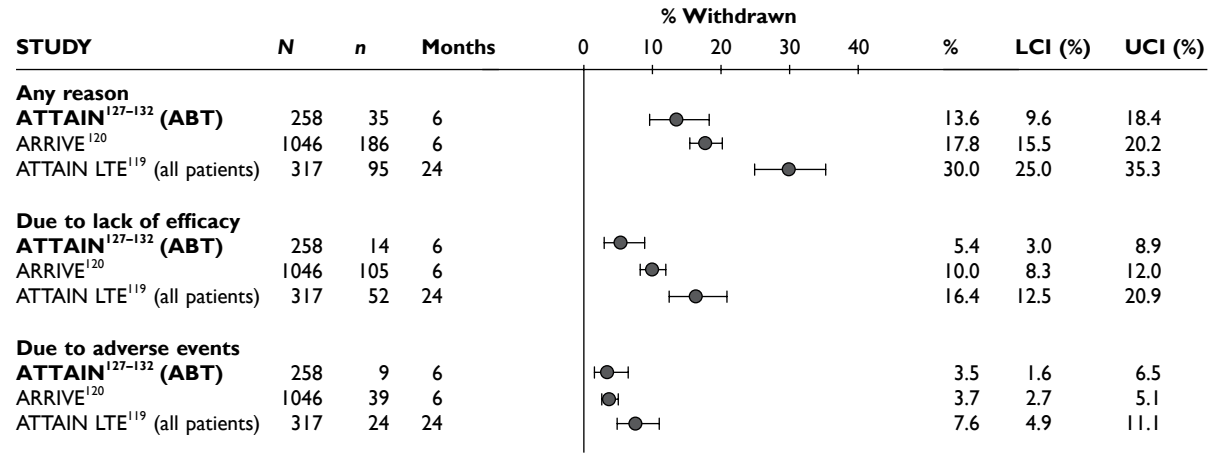


FIGURE 59 Abatacept: withdrawals in uncontrolled studies by reason. LCI, lower confidence interval; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

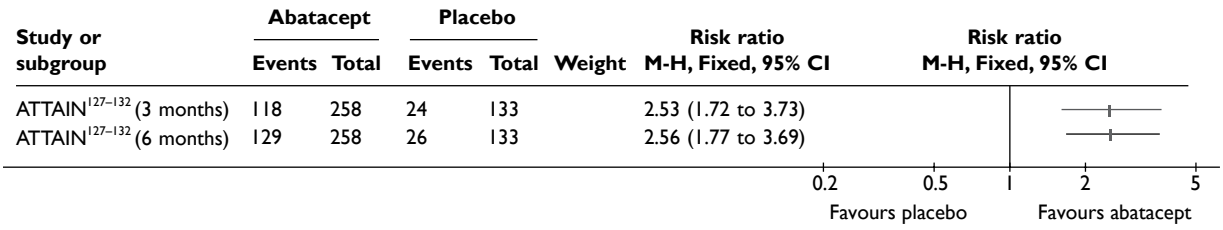
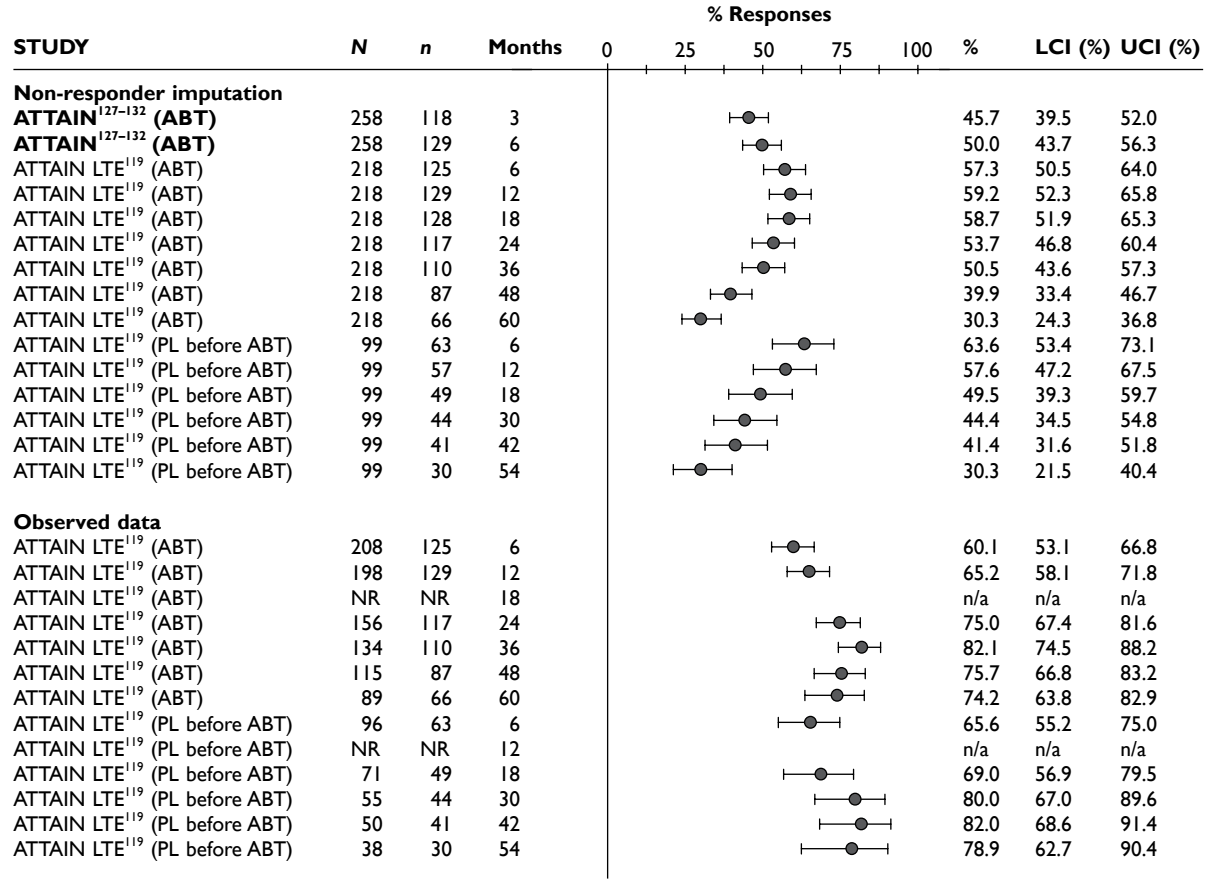


FIGURE 60 Abatacept: ACR20 response in the ATTAIN RCT[127–132] at 3 and 6 months.

randomised to ABT or placebo in the randomised phase (see *Figure 63*). This outcome was achieved at 6 months by 22.9% patients in the arm initially randomised to ABT and 37.4% in the arm initially randomised to placebo. For comparison, this outcome was achieved by 20.2% of patients in the ABT arm of the RCT. In the arm initially randomised to ABT, the percentage of ACR50 responders increased up to 18 months (33.9%) and then decreased to 20.6% at 5 years. In the arm initially randomised to placebo, there was a decrease after 6 months to 21.2% achieving ACR50 response at 48 months.

In the analysis based on the observed data, only the percentage of ACR50 responders among those initially randomised to ABT increased up to 3 years (51.1%) and then it was 46.1% at



STUDY	N	n	Months	%	LCI (%)	UCI (%)
Non-responder imputation						
ATTAIN[127–132] (ABT)	258	118	3	45.7	39.5	52.0
ATTAIN[127–132] (ABT)	258	129	6	50.0	43.7	56.3
ATTAIN LTE[119] (ABT)	218	125	6	57.3	50.5	64.0
ATTAIN LTE[119] (ABT)	218	129	12	59.2	52.3	65.8
ATTAIN LTE[119] (ABT)	218	128	18	58.7	51.9	65.3
ATTAIN LTE[119] (ABT)	218	117	24	53.7	46.8	60.4
ATTAIN LTE[119] (ABT)	218	110	36	50.5	43.6	57.3
ATTAIN LTE[119] (ABT)	218	87	48	39.9	33.4	46.7
ATTAIN LTE[119] (ABT)	218	66	60	30.3	24.3	36.8
ATTAIN LTE[119] (PL before ABT)	99	63	6	63.6	53.4	73.1
ATTAIN LTE[119] (PL before ABT)	99	57	12	57.6	47.2	67.5
ATTAIN LTE[119] (PL before ABT)	99	49	18	49.5	39.3	59.7
ATTAIN LTE[119] (PL before ABT)	99	44	30	44.4	34.5	54.8
ATTAIN LTE[119] (PL before ABT)	99	41	42	41.4	31.6	51.8
ATTAIN LTE[119] (PL before ABT)	99	30	54	30.3	21.5	40.4
Observed data						
ATTAIN LTE[119] (ABT)	208	125	6	60.1	53.1	66.8
ATTAIN LTE[119] (ABT)	198	129	12	65.2	58.1	71.8
ATTAIN LTE[119] (ABT)	NR	NR	18	n/a	n/a	n/a
ATTAIN LTE[119] (ABT)	156	117	24	75.0	67.4	81.6
ATTAIN LTE[119] (ABT)	134	110	36	82.1	74.5	88.2
ATTAIN LTE[119] (ABT)	115	87	48	75.7	66.8	83.2
ATTAIN LTE[119] (ABT)	89	66	60	74.2	63.8	82.9
ATTAIN LTE[119] (PL before ABT)	96	63	6	65.6	55.2	75.0
ATTAIN LTE[119] (PL before ABT)	NR	NR	12	n/a	n/a	n/a
ATTAIN LTE[119] (PL before ABT)	71	49	18	69.0	56.9	79.5
ATTAIN LTE[119] (PL before ABT)	55	44	30	80.0	67.0	89.6
ATTAIN LTE[119] (PL before ABT)	50	41	42	82.0	68.6	91.4
ATTAIN LTE[119] (PL before ABT)	38	30	54	78.9	62.7	90.4

FIGURE 61 Abatacept: ACR20 response in non-RCTs. LCI, lower confidence interval; n/a, not available; PL, placebo; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

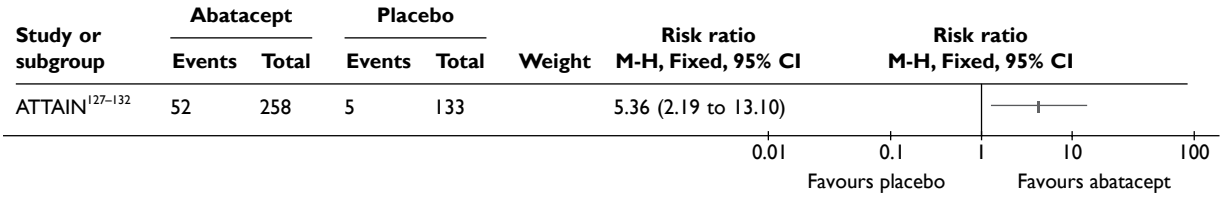


FIGURE 62 Abatacept: ACR50 response in the ATTAIN RCT[127–132] at 6 months.

4 years and 51.1% at 5 years. Among those initially randomised to placebo there was an almost constant increase up to 48 months (53.8%).

ACR 70 response

Randomised controlled trial In the ATTAIN RCT,[127–132] the percentage of patients achieving ACR70 response at 6 months was almost seven times higher in the ABT group than in the placebo group (RR = 6.70, 95% CI 1.62 to 27.8). This difference was statistically significant; however, it needs to be highlighted that the CIs were very wide (see *Figure 64*).

Non-randomised controlled trials Of the uncontrolled studies, only the ATTAIN LTE[119] reported ACR70 response. After 6 months of treatment the percentage of ACR70 responders was 11.5% among patients initially treated with ABT and 13.1% among patients initially treated with placebo. For comparison, it was 10.1% in the ATTAIN RCT.[127–132] In the arm initially randomised

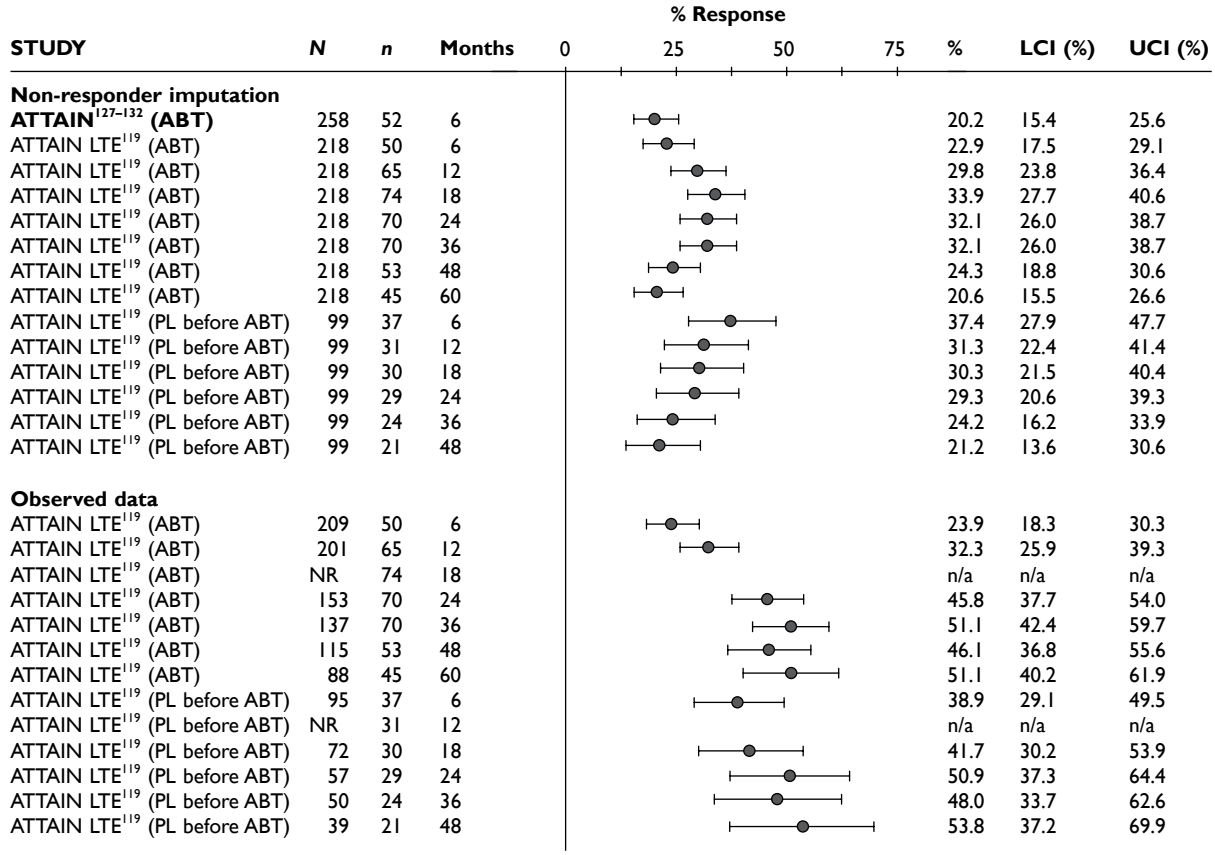


FIGURE 63 Abatacept: ACR50 response in non-RCTs. LCI, lower confidence interval; n/a, not available; PL, placebo; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

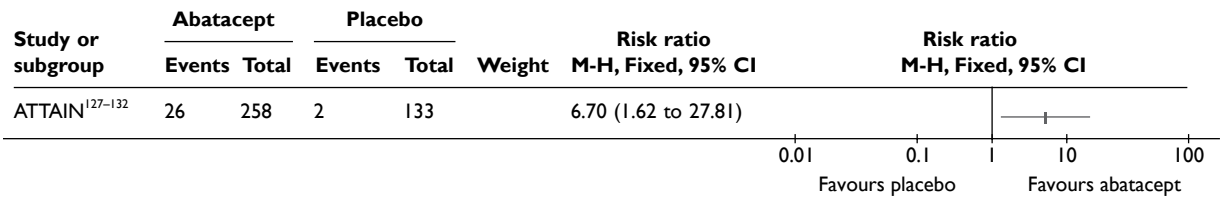


FIGURE 64 Abatacept: ACR70 response in the ATTAIN RCT[127–132] at 6 months.

to ABT, there was a further increase to 17.0% at 12 months followed by a decrease to 9.6% at 5 years. In the arm initially randomised to placebo, there was an increase up to 15.2% at 30 months followed by a decrease to 7.1% at 54 months. Analysis based on observed data only provided more favourable results, with the highest percentage of ACR70 responders being 23.4% at 36 months in the arm initially randomised to ABT and 25.9% at 30 months in the arm initially randomised to placebo. See *Figure 65* for details.

DAS28

Randomised controlled trial The mean change from baseline in DAS28 was –1.98 in the ABT group and –0.71 in the placebo group. The difference between these values was –1.27 (95% CI –1.62 to –0.93, $p < 0.001$). These data were provided in the industry submission only. No further information was provided and therefore analyses could not be undertaken.

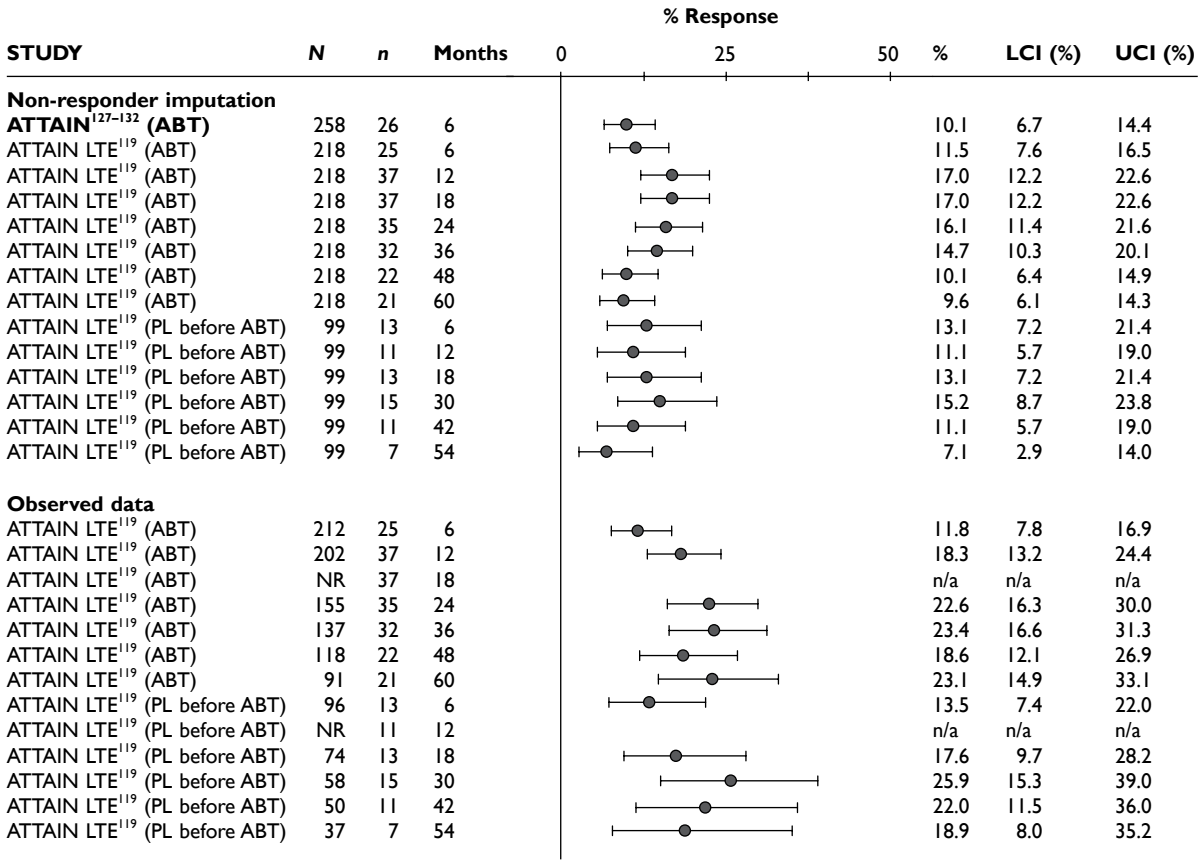


FIGURE 65 Abatacept: ACR70 response in non-RCTs. LCI, lower confidence interval; n/a, not available; PL, placebo; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

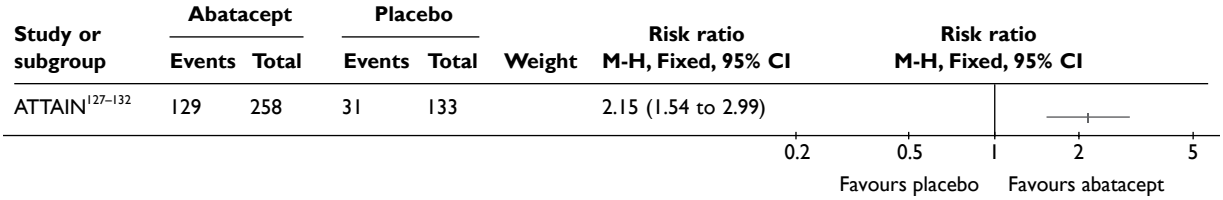


FIGURE 66 Abatacept: patients with clinically meaningful (≥ 1.2) DAS28 improvement in the ATTAIN RCT[127–132] at 6 months.

As indicated in *Figure 66*, over twice as many patients achieved a clinically meaningful DAS28 improvement (defined as greater than or equal to 1.2) in the ABT arm as in the control arm (RR = 2.15, 95% CI 1.54 to 2.99).

The ATTAIN study[127–132] also reported percentages of patients who, based on DAS28, achieved a low score (DAS28 less than or equal to 3.2) or remission (DAS28 less than 2.6). At 6 months, patients in the ABT arm were over five times more likely to have a DAS28 less than or equal to 3.2 than those in the placebo arm and the difference was statistically significant (RR = 5.67, 95% CI 2.08 to 15.44). They were also over 13 times more likely to have a DAS28 less than 2.6 than the placebo group and the difference was statistically significant (RR = 13.40, 95% CI 1.84 to 97.69); however, the CIs were wide. See *Figure 67* for details.

Non-randomised controlled trials Change in the DAS28 score was assessed in both uncontrolled studies. Details are presented in *Figure 68*. After 6 months of treatment, there was a mean change of –1.99 in the arm initially randomised to ABT and of –2.14 in the arm initially randomised to placebo in the ATTAIN LTE,[119] and of –2.00 in the ARRIVE study.[120] This was similar in the RCT.[127–132] In the ATTAIN LTE,[119] DAS28 further decreased with time and the mean change was –2.90 at 5 years in the arm initially randomised to ABT and –2.96 at 54 months in the arm initially randomised to placebo.

ARRIVE[120] measured clinically meaningful DAS28 improvement. It was defined as a decrease of greater than or equal to 1.2 or a score of less than or equal to 3.2. At 6 months, 56.1% of patients in ARRIVE[120] achieved this outcome. This was slightly more than in the ABT group of the RCT[127–132] (although in ATTAIN[127–132] this was defined as a decrease of greater than or equal to 1.2 only). See *Figure 69* for details.

Both uncontrolled studies reported percentages of patients who, based on DAS28, achieved a low score (DAS28 less than or equal to 3.2) or remission (DAS28 less than 2.6). Full details are reported in *Figure 70*.

Study or subgroup	Abatacept Events	Abatacept Total	Placebo Events	Placebo Total	Weight	Risk ratio M-H, Fixed, 95% CI
DAS28 ≤ 3.2						
ATTAIN[127–132]	44	258	4	133		5.67 (2.08 to 15.44)
DAS28 < 2.6						
ATTAIN[127–132]	26	258	1	133		13.40 (1.84 to 97.69)

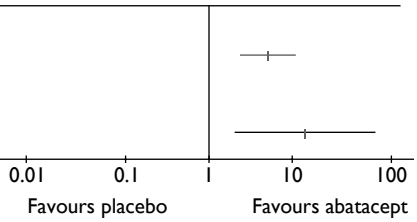


FIGURE 67 Abatacept: patients with final DAS28 of ≤3.2 and of < 2.6 in the ATTAIN RCT[127–132] at 6 months.

STUDY	*N*	Mean	SD	Months	95% LCI	95% UCI
ATTAIN[127–132] (ABT)	258	–1.98	NR	6	n/a	n/a
ARRIVE[120]	1046	–2.00	2.32	6	–2.14	–1.86
ATTAIN LTE[119] (ABT)	205	–1.99	NR	6	–2.19	–1.80
ATTAIN LTE[119] (ABT)	192	–2.33	NR	12	–2.42	–2.13
ATTAIN LTE[119] (ABT)	151	–2.66	NR	24	–2.87	–2.44
ATTAIN LTE[119] (ABT)	132	–2.85	NR	36	–3.09	–2.62
ATTAIN LTE[119] (ABT)	113	–2.79	NR	48	–3.04	–2.54
ATTAIN LTE[119] (ABT)	79	–2.90	NR	60	–3.22	–2.58
ATTAIN LTE[119] (PL before ABT)	93	–2.14	NR	6	–2.43	–1.84
ATTAIN LTE[119] (PL before ABT)	68	–2.23	NR	18	–2.56	–1.90
ATTAIN LTE[119] (PL before ABT)	56	–2.62	NR	30	–3.02	–2.22
ATTAIN LTE[119] (PL before ABT)	49	–2.77	NR	42	–3.16	–2.38
ATTAIN LTE[119] (PL before ABT)	33	–2.96	NR	54	–3.34	–2.58

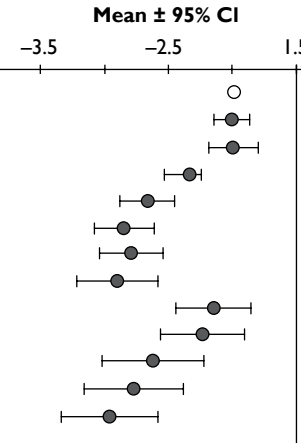


FIGURE 68 Abatacept: DAS28 change from baseline in uncontrolled studies. LCI, lower confidence interval; n/a, not available; PL, placebo; SD, standard deviation; UCI, upper confidence interval.

STUDY	*N*	*n*	Months	%	LCI (%)	UCI (%)
ATTAIN[127–132] (ABT)	258	129	6	50.0	43.7	56.3
ARRIVE[120]	1046	587	6	56.1	53.0	59.2

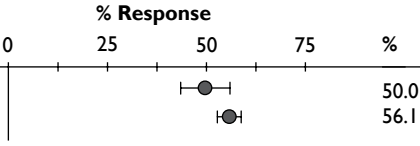


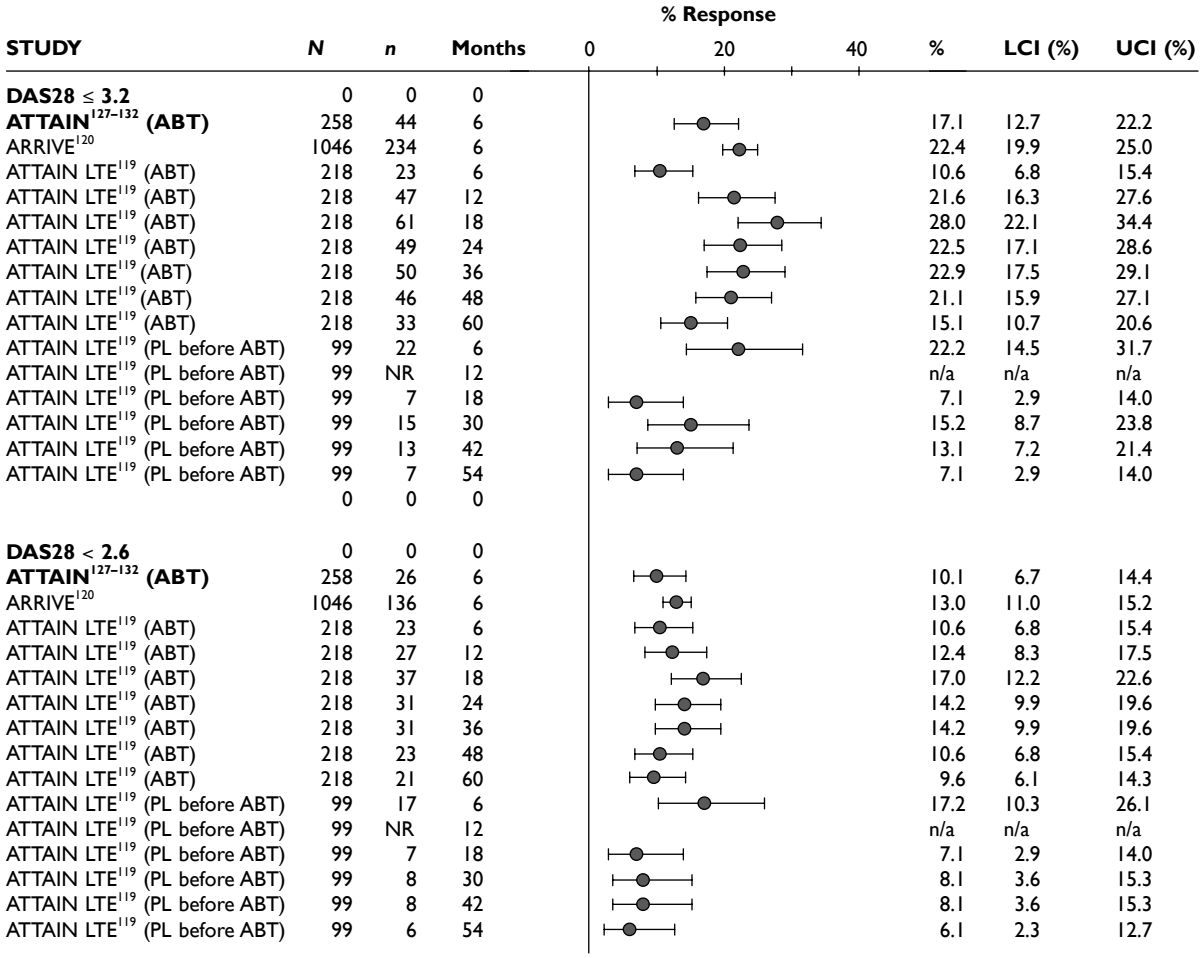
FIGURE 69 Abatacept: clinically meaningful DAS28 improvement in non-randomised studies at 6 months.

At 6 months a DAS28 score of less than or equal to 3.2 was achieved by 10.6% of patients initially randomised to ABT in the ATTAIN LTE,[119] by 22.2% of patients initially randomised to placebo in the ATTAIN LTE[119] and by 22.4% of patients in ARRIVE.[120] For comparison, this was 17.1% of patients in the ABT arm of ATTAIN.[127–132] The percentage of patients initially randomised to ABT in ATTAIN LTE[119] who achieved a DAS28 of less than or equal to 3.2 increased up to 18 months (28%) and then decreased up to 5 years (15.1%). In the arm initially randomised to placebo, the percentage of patients with low DAS28 decreased up to 54 months (7.1%).

A DAS28 of less than 2.6 was achieved at 6 months by 10.6% and 17.2% in the ATTAIN LTE[119] (initial ABT and placebo, respectively) and by 13.0% in ARRIVE.[120] For comparison, 10.1% of the ABT arm of the RCT achieved this outcome. In the ATTAIN LTE[119] arm initially randomised to ABT, the highest percentage of patients with DAS28 less than 2.6 was recorded at 18 months (17.0%), following which it decreased to 9.6% at 5 years. In the arm initially randomised to placebo, the highest percentage of patients with DAS28 less than 2.6 was recorded after 6 months of treatment, and at 54 months it was 6.1%.

EULAR response

EULAR response was not assessed in any of the studies.



STUDY	N	n	Months	%	LCI (%)	UCI (%)
DAS28 ≤ 3.2	0	0	0			
ATTAIN[127–132] (ABT)	258	44	6	17.1	12.7	22.2
ARRIVE[120]	1046	234	6	22.4	19.9	25.0
ATTAIN LTE[119] (ABT)	218	23	6	10.6	6.8	15.4
ATTAIN LTE[119] (ABT)	218	47	12	21.6	16.3	27.6
ATTAIN LTE[119] (ABT)	218	61	18	28.0	22.1	34.4
ATTAIN LTE[119] (ABT)	218	49	24	22.5	17.1	28.6
ATTAIN LTE[119] (ABT)	218	50	36	22.9	17.5	29.1
ATTAIN LTE[119] (ABT)	218	46	48	21.1	15.9	27.1
ATTAIN LTE[119] (ABT)	218	33	60	15.1	10.7	20.6
ATTAIN LTE[119] (PL before ABT)	99	22	6	22.2	14.5	31.7
ATTAIN LTE[119] (PL before ABT)	99	NR	12	n/a	n/a	n/a
ATTAIN LTE[119] (PL before ABT)	99	7	18	7.1	2.9	14.0
ATTAIN LTE[119] (PL before ABT)	99	15	30	15.2	8.7	23.8
ATTAIN LTE[119] (PL before ABT)	99	13	42	13.1	7.2	21.4
ATTAIN LTE[119] (PL before ABT)	99	7	54	7.1	2.9	14.0
	0	0	0			
DAS28 < 2.6	0	0	0			
ATTAIN[127–132] (ABT)	258	26	6	10.1	6.7	14.4
ARRIVE[120]	1046	136	6	13.0	11.0	15.2
ATTAIN LTE[119] (ABT)	218	23	6	10.6	6.8	15.4
ATTAIN LTE[119] (ABT)	218	27	12	12.4	8.3	17.5
ATTAIN LTE[119] (ABT)	218	37	18	17.0	12.2	22.6
ATTAIN LTE[119] (ABT)	218	31	24	14.2	9.9	19.6
ATTAIN LTE[119] (ABT)	218	31	36	14.2	9.9	19.6
ATTAIN LTE[119] (ABT)	218	23	48	10.6	6.8	15.4
ATTAIN LTE[119] (ABT)	218	21	60	9.6	6.1	14.3
ATTAIN LTE[119] (PL before ABT)	99	17	6	17.2	10.3	26.1
ATTAIN LTE[119] (PL before ABT)	99	NR	12	n/a	n/a	n/a
ATTAIN LTE[119] (PL before ABT)	99	7	18	7.1	2.9	14.0
ATTAIN LTE[119] (PL before ABT)	99	8	30	8.1	3.6	15.3
ATTAIN LTE[119] (PL before ABT)	99	8	42	8.1	3.6	15.3
ATTAIN LTE[119] (PL before ABT)	99	6	54	6.1	2.3	12.7

FIGURE 70 Abatacept: patients with final DAS28 values ≤3.2 and <2.6 in uncontrolled studies. LCI, lower confidence interval; n/a, not available; PL, placebo; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

Health Assessment Questionnaire

Randomised controlled trial At 6 months, the HAQ change from baseline in the ATTAIN RCT[127–132] was –0.45 in the ABT group and –0.11 in the placebo group and the difference between the two groups was reported to be statistically significant ($p < 0.001$). No data on uncertainty of individual assessments were provided in the study and therefore further analyses could not be undertaken.

This study also assessed clinically meaningful HAQ improvement, defined as a decrease in HAQ score of at least 0.3 (details are reported in *Figure 71*). Clinically meaningful HAQ improvement was over two times more frequent in the ABT group than in the placebo group and the difference was statistically significant (RR = 2.01, 95% CI 1.44 to 2.81).

Non-randomised control trials Change in HAQ score was assessed in both uncontrolled studies (however, in the case of for ARRIVE[120] only data for a subgroup of 43 US patients receiving monotherapy were reported; ABT monotherapy is licensed in the USA but not in Europe). *Figure 72* presents the mean changes from baseline in HAQ score. The mean change from baseline at 6 months was –0.51 in the arm of ATTAIN[127–132] initially randomised to ABT, –0.40 in the arm of ATTAIN[127–132] initially randomised to placebo and –0.38 in the monotherapy subgroup of ARRIVE.[120] The results for the ABT arm of the RCT were similar. In the arm initially randomised to ABT in the ATTAIN LTE,[119] the change decreased up to 3 years (–0.65) and then started slowly increasing (to –0.58 at 4 years and to –0.56 at 5 years). In the group initially randomised to placebo, there was a decrease up to 54 months of treatment (–0.71).

Both uncontrolled studies reported the number of patients who achieved a clinically meaningful improvement in HAQ (details are provided in *Figure 73*). The ATTAIN LTE[119] defined this outcome as an improvement of at least 0.3 in the HAQ score, while in ARRIVE[120] it was an improvement of at least 0.22. After 6 months of treatment with ABT, the percentage of

Study or subgroup	Abatacept Events	Abatacept Total	Placebo Events	Placebo Total	Weight	Risk ratio M-H, Fixed, 95% CI
ATTAIN[127–132]	121	258	31	133		2.01 (1.44 to 2.81)

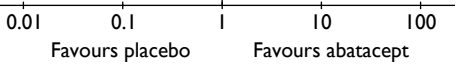


FIGURE 71 Abatacept: clinically meaningful improvement (≥0.3) in HAQ score.

STUDY	N	Mean	SD	Months	95% LCI	95% UCI
ATTAIN[127–132] (ABT)	258	–0.45	NR	6	NR	NR
ARRIVE[120] (montherapy)	43	–0.38	0.66	6	–0.58	–0.18
ATTAIN LTE[119] (ABT)	207	–0.51	NR	6	–0.59	–0.43
ATTAIN LTE[119] (ABT)	199	–0.52	NR	12	–0.60	–0.45
ATTAIN LTE[119] (ABT)	154	–0.62	NR	24	–0.71	–0.52
ATTAIN LTE[119] (ABT)	128	–0.65	NR	36	–0.74	–0.55
ATTAIN LTE[119] (ABT)	117	–0.58	NR	48	–0.67	–0.48
ATTAIN LTE[119] (ABT)	87	–0.56	NR	60	–0.69	–0.43
ATTAIN LTE[119] (PL before ABT)	97	–0.40	NR	6	–0.50	–0.30
ATTAIN LTE[119] (PL before ABT)	74	–0.50	NR	18	–0.64	–0.37
ATTAIN LTE[119] (PL before ABT)	54	–0.51	NR	30	–0.66	–0.37
ATTAIN LTE[119] (PL before ABT)	50	–0.62	NR	42	–0.80	–0.43
ATTAIN LTE[119] (PL before ABT)	40	–0.71	NR	54	–0.90	–0.51

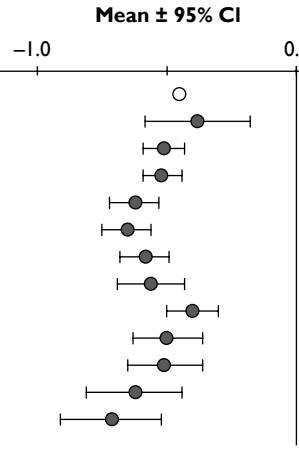
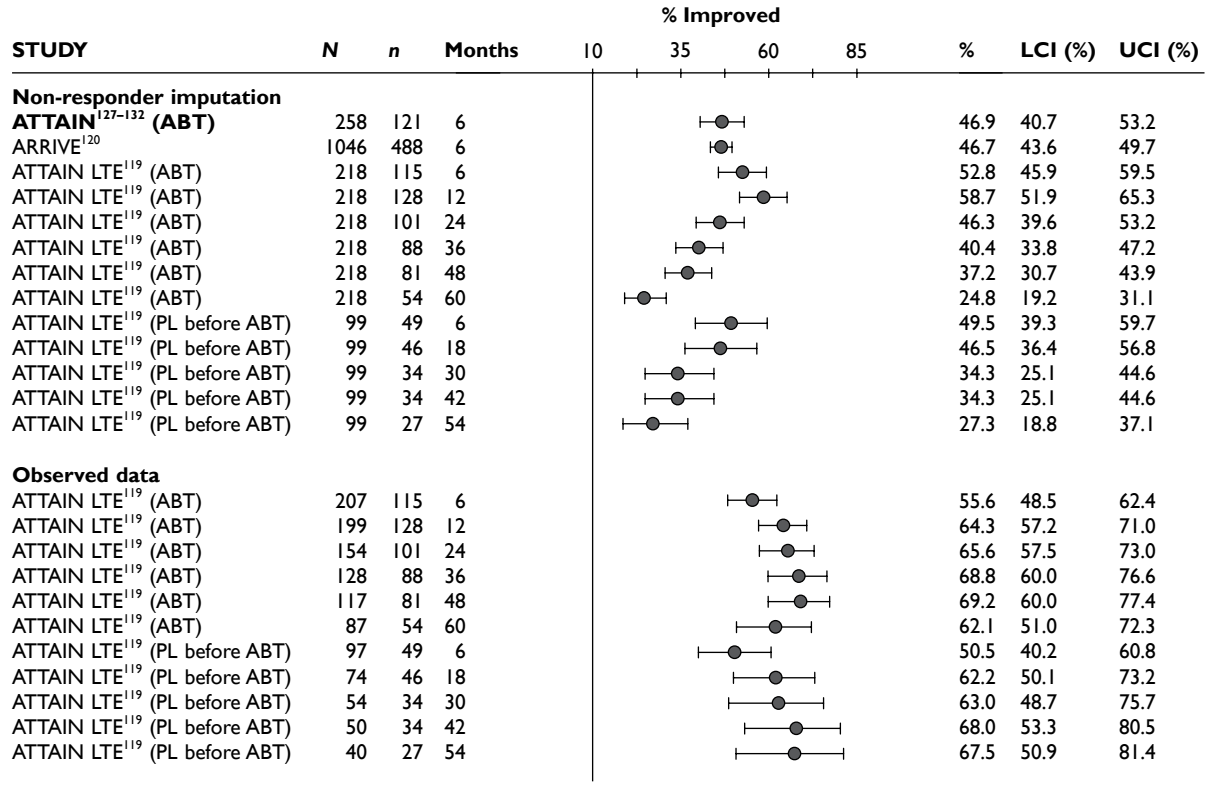


FIGURE 72 Abatacept: mean changes from baseline in HAQ score. LCI, lower confidence interval; PL, placebo; SD, standard deviation; UCI, upper confidence interval.

patients who achieved this outcome was 52.8% in the ATTAIN LTE[119] arm that comprised patients initially randomised to ABT, 49.5% in the ATTAIN LTE[119] arm comprising patients initially randomised to placebo and 46.7% in the ARRIVE study.[120] For comparison, it was 46.9% in the ABT arm of the RCT. Analysis of the data from the ATTAIN LTE[119] using a non-responder imputation showed a decrease in the percentage of patients who achieved a clinically meaningful HAQ over time, with 24.8% of patients initially randomised to ABT achieving clinically meaningful HAQ improvement at 5 years and 27.3% of patients initially randomised to placebo achieving clinically meaningful HAQ improvement at 54 months. When the analysis in both groups included only patients in whom HAQ improvement was measured at different time points, there was a slight increase in the percentage over time, with a decrease in the last outcome measurement.



STUDY	N	n	Months	%	LCI (%)	UCI (%)
Non-responder imputation						
ATTAIN[127–132] (ABT)	258	121	6	46.9	40.7	53.2
ARRIVE[120]	1046	488	6	46.7	43.6	49.7
ATTAIN LTE[119] (ABT)	218	115	6	52.8	45.9	59.5
ATTAIN LTE[119] (ABT)	218	128	12	58.7	51.9	65.3
ATTAIN LTE[119] (ABT)	218	101	24	46.3	39.6	53.2
ATTAIN LTE[119] (ABT)	218	88	36	40.4	33.8	47.2
ATTAIN LTE[119] (ABT)	218	81	48	37.2	30.7	43.9
ATTAIN LTE[119] (ABT)	218	54	60	24.8	19.2	31.1
ATTAIN LTE[119] (PL before ABT)	99	49	6	49.5	39.3	59.7
ATTAIN LTE[119] (PL before ABT)	99	46	18	46.5	36.4	56.8
ATTAIN LTE[119] (PL before ABT)	99	34	30	34.3	25.1	44.6
ATTAIN LTE[119] (PL before ABT)	99	34	42	34.3	25.1	44.6
ATTAIN LTE[119] (PL before ABT)	99	27	54	27.3	18.8	37.1
Observed data						
ATTAIN LTE[119] (ABT)	207	115	6	55.6	48.5	62.4
ATTAIN LTE[119] (ABT)	199	128	12	64.3	57.2	71.0
ATTAIN LTE[119] (ABT)	154	101	24	65.6	57.5	73.0
ATTAIN LTE[119] (ABT)	128	88	36	68.8	60.0	76.6
ATTAIN LTE[119] (ABT)	117	81	48	69.2	60.0	77.4
ATTAIN LTE[119] (ABT)	87	54	60	62.1	51.0	72.3
ATTAIN LTE[119] (PL before ABT)	97	49	6	50.5	40.2	60.8
ATTAIN LTE[119] (PL before ABT)	74	46	18	62.2	50.1	73.2
ATTAIN LTE[119] (PL before ABT)	54	34	30	63.0	48.7	75.7
ATTAIN LTE[119] (PL before ABT)	50	34	42	68.0	53.3	80.5
ATTAIN LTE[119] (PL before ABT)	40	27	54	67.5	50.9	81.4

FIGURE 73 Abatacept: clinically meaningful improvement in HAQ score (≥0.3 in ATTAIN studies[119,127–132] and ≥0.22 in ARRIVE[120]) LCI, lower confidence interval; PL, placebo; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

Study or subgroup	Abatacept Mean	Abatacept SD	Abatacept Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean difference IV, Fixed, 95% CI
Physical component								
ATTAIN[127–132]	6.5	9.6	258	1	7.7	133		5.50 (3.74 to 7.26)
Mental component								
ATTAIN[127–132]	5.4	11.7	258	1.7	10.2	133		3.70 (1.45 to 5.95)

FIGURE 74 Abatacept: SF-36 items changes from baseline in components in the ATTAIN RCT[127–132] at 6 months. SD, standard deviation.

Quality of life

Randomised controlled trial The ATTAIN RCT[127–132] assessed patients' QoL using the SF-36 scale. Patients in the ABT arm improved significantly more in both the physical component (mean difference = 5.50, 95% CI 3.74 to 7.26) and the mental component (mean difference = 3.70, 95% CI 1.45 to 5.95). Details are presented in *Figure 74*.

For all individual SF-36 items there was a significantly higher improvement in the ABT arm than in the placebo arm. Details for each item are presented in *Figure 75*.

Non-randomised controlled trials Of the uncontrolled studies, change in SF-36 was assessed only in the ARRIVE study[120] (however, it was reported only for a subgroup of 43 patients receiving monotherapy; ABT monotherapy is licensed in the USA but not in Europe). For the physical component of the SF-36 scale, there was improvement of 4.80 for the monotherapy subgroup of ARRIVE.[120] For the mental component, the improvement was 7.34. For comparison, in the ABT arm of ATTAIN[127–132] it was 6.50 and 5.40, respectively. Further details are provided in *Figure 76*. Data for individual items were not reported in ARRIVE.[120]

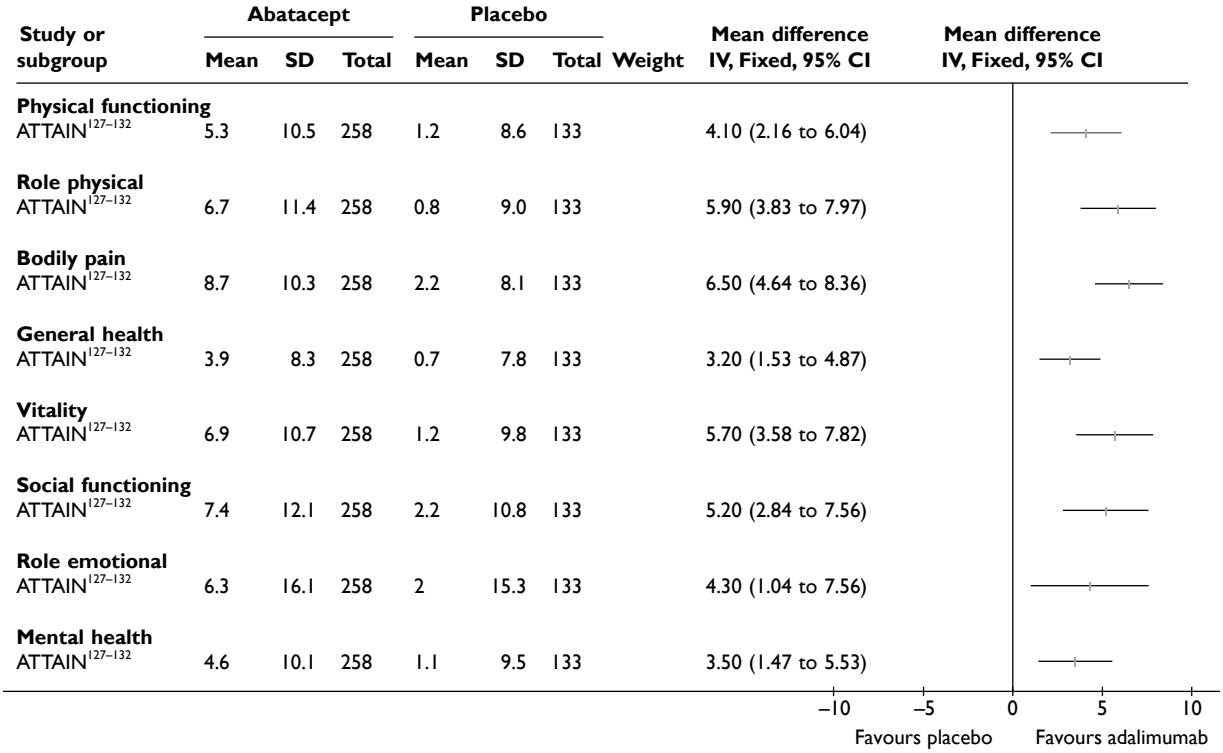


FIGURE 75 Abatacept: SF-36 items changes from baseline in items at 6 months in the ATTAIN RCT.[127–132] SD, standard deviation.

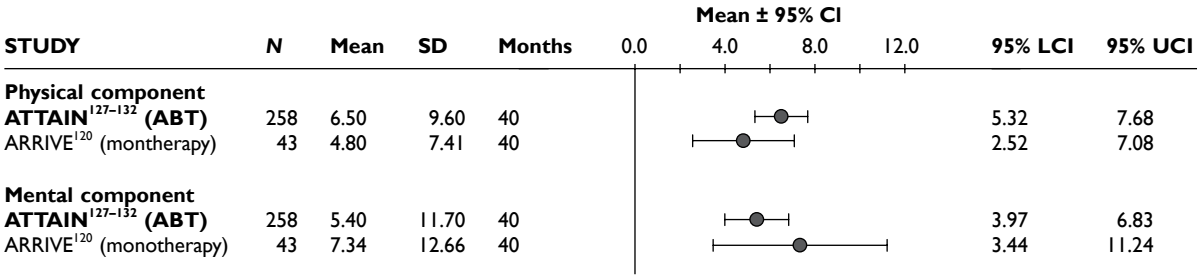


FIGURE 76 Abatacept: SF-36 items changes from baseline in components. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

Joint damage

Joint damage was not assessed in any of the studies.

Serious adverse events

Randomised controlled trial In ATTAIN,[127–132] there was no significant difference at 6 months between ABT and placebo in the risk of experiencing a serious AE (RR = 0.93, 95% CI 0.51 to 1.68). Details are presented in *Figure 77*.

Non-randomised controlled trials Serious AEs were assessed in both uncontrolled studies. At 6 months the percentage of patients who had experienced a serious AE was 10.4% in ARRIVE.[120] It was similar in the ABT arm of the ATTAIN RCT[127–132] (10.5%). At 2 years, 32.5% of patients in the ATTAIN LTE[119] had experienced a serious AE. Full details are presented in *Figure 78*.

Infections/serious infections

Randomised controlled trial At 6 months there was no statistically significant difference between ABT and placebo in the risk of infection (RR = 1.16, 95% CI 0.87 to 1.56) or serious infection (RR = 1.03, 95% CI 0.26 to 4.06). Details are presented in *Figure 79*.

Non-randomised controlled trials Both uncontrolled studies reported infections. The percentages of patients who experienced any infection were similar at 6 months in the ABT arm of ATTAIN[127–132] and in the ARRIVE study[120] (37.6% and 38.9%, respectively). Of these 2.3% and 2.4% were serious. At 2 years 73.8% of patients in the ATTAIN LTE[119] experienced an infection of any kind and 7.9% a serious infection. Details are reported in *Figure 80*.

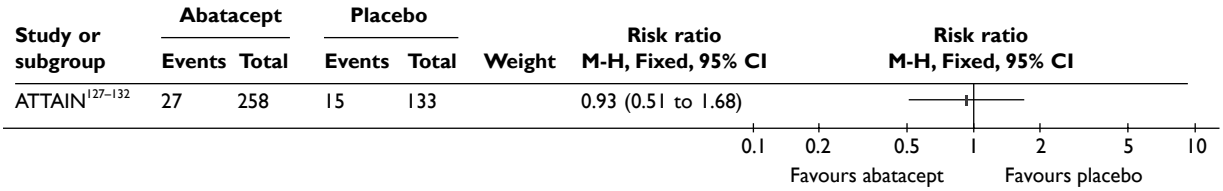


FIGURE 77 Abatacept: serious AEs in the ATTAIN RCT[127–132] at 6 months.

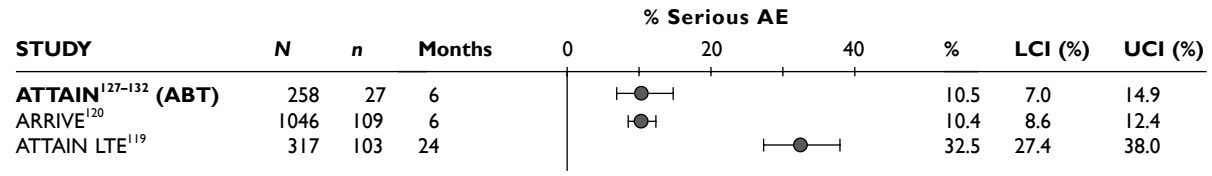


FIGURE 78 Abatacept: serious adverse events in non-randomised studies. LCI, lower confidence interval; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

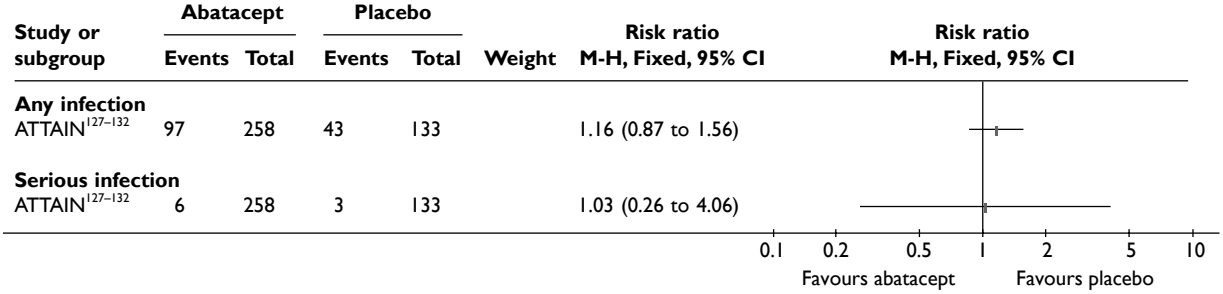


FIGURE 79 Abatacept: infections in the ATTAIN RCT[127–132] at 6 months.

Injection/infusion reaction

Randomised controlled trial At 6 months there was no statistically significant difference between ABT and placebo in the risk of infusion reaction (RR = 1.68, 95% CI 0.56 to 5.04). Details are reported in *Figure 81*.

Non-randomised controlled trials Of the uncontrolled studies, infusion reactions were reported only in ARRIVE.[120] At 6 months, 5.4% patients had experienced infusion reactions. For comparison, this figure was 5.0% in the ABT arm of ATTAIN.[127–132] Details are provided in *Figure 82*.

Abatacept in combination with other biologic drugs

Two RCTs [Weinblatt *et al.*[134] and abatacept study of safety in use with other rheumatoid arthritis therapies (ASSURE)[135]] were identified that assessed ABT in combination with previously tried biologic drugs. Although both studies met the inclusion criteria of the systematic review, combination therapy was not considered relevant to this report and, therefore, they were not analysed.

The study by Weinblatt *et al.*[134] was a multicentre placebo-controlled randomised trial and included 121 patients who had active RA despite treatment with ETN. Patients were randomised to receive ETN and ABT or ETN and placebo and were followed up for 1 year. Afterwards they could enter a LTE (data provided for 2 years of the extension study). Data were collected on outcomes including ACR response, HAQ, SF-36 and safety.

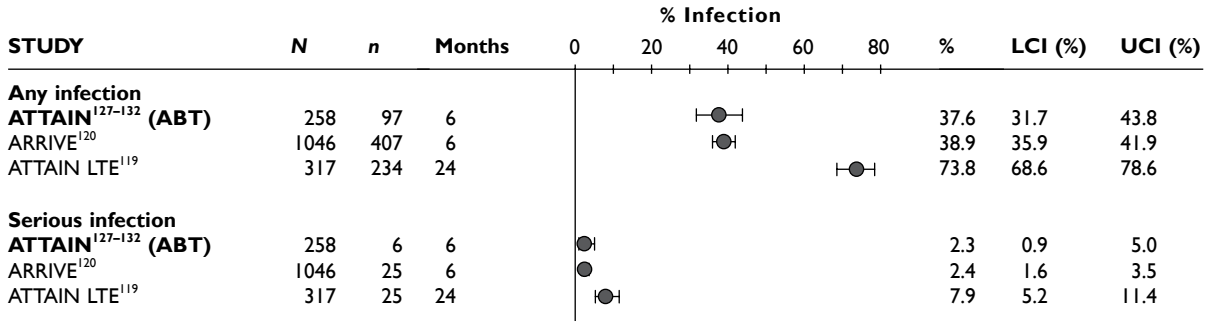


FIGURE 80 Abatacept: infections in non-randomised studies.

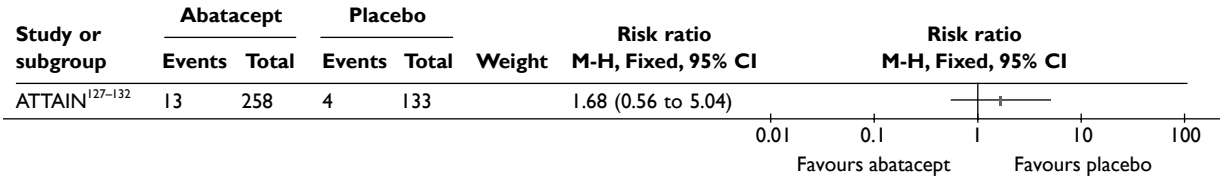


FIGURE 81 Abatacept: infusion reactions. LCI, lower confidence interval; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

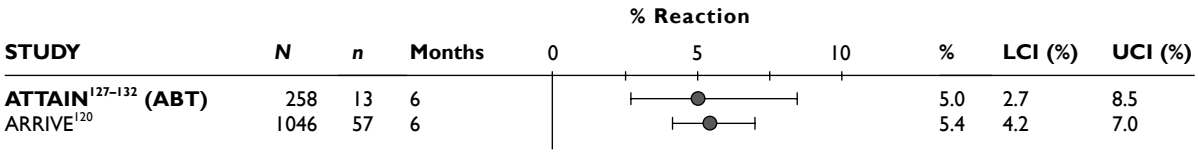


FIGURE 82 Abatacept: infusion reactions. LCI, lower confidence interval; UCI, upper confidence interval. **Bold type** indicates that the study was an RCT.

ASSURE[135] was a multicentre placebo-controlled randomised trial and included 167 patients who had active RA in spite of receiving therapy with biologic agents (ETN, IFX, ADA and anakinra), 'warranting additional therapy at the discretion of the investigator'. (Note: it also included 1,274 patients who received background DMARDs and were probably biologic naive.) Patients continued their treatment and in addition to that were randomised to receive ABT or placebo. They were followed up for 1 year. The study assessed outcomes including HAQ Disability Index, pain, patient and physician global assessment and safety.

Summary

Three studies assessed ABT in comparison with standard care: one RCT (ATTAIN[127–132]) and two uncontrolled studies (ATTAIN LTE[119] and ARRIVE[120]). Follow-up ranged from 6 months to 5 years. All studies included patients with similar baseline characteristics. The main results of the included studies are summarised in *Table 40*.

TABLE 40 Abatacept: summary of main results

	RCT [result (95% CI)]		Uncontrolled studies	
Outcome	6 months (ABT vs placebo)	6 months (ABT arm)	6 months	4.5–5 years
Withdrawals:				24 months (longer follow-up NA)
■ for any reason	RR = 0.53 (0.35 to 0.81); less in ABT	13.6%	17.8%	30%
■ due to lack of efficacy	RR = 0.27 (0.15 to 0.49), less in ABT	5.4%	10%	16.4%
■ due to AEs	RR = 0.93 (0.32 to 2.71), no difference	3.5%	3.7%	7.6%
ACR20 response	RR = 2.56 (1.77 to 3.69), favours ABT; similar results for 3 months	50.0%	57.3%–63.6%	30.3%
ACR50 response	RR = 5.36 (2.19 to 13.10), favours ABT	20.2%	22.9%–37.4%	20.6%–21.2%
ACR70 response	RR = 6.70 (1.62 to 27.81), favours ABT	10.1%	11.5%–13.1%	7.1%–9.6%
DAS28:				
■ change from baseline	Mean difference = –1.27 (–1.62 to –0.93), favours ABT	–1.98	–1.99 to –2.14	–2.00 to –2.90
■ clinically meaningful		50.0%	56.1%	NA
■ ≤ 3.2	RR = 2.15 (1.54 to 2.99), favours ABT	17.1%	10.6%–22.4%	7.1%–15.1%
■ < 2.6	RR = 5.67 (2.08 to 15.44), favours ABT	10.1%	13.0%–17.2%	6.1%–9.6%
	RR = 13.40 (1.84 to 97.69), favours ABT			
EULAR response	NA	NA	NA	NA
HAQ:				
■ change from baseline	Mean difference = –0.34, favours ABT ($p < 0.001$)	–0.45	–0.38 to –0.51	–0.56 to –0.71
■ clinically meaningful		46.9%	46.7%–52.8%	24.8%–27.3%
	RR = 2.01 (1.44 to 2.81), favours ABT			
QoL (SF-36)				
■ physical component, change from baseline	Mean difference = 5.50 (3.74 to 7.26), favours ABT	6.50	7.41	NA
■ mental component, change from baseline	Mean difference = 3.70 (1.45 to 5.95), favours ABT	5.40	12.66	NA
Joint damage	NA	NA	NA	NA
Serious AEs	RR = 0.93 (0.51 to 1.68), NS	10.5%	10.4%	32.5%
Any infections	RR = 1.16 (0.87 to 1.56), NS	37.6%	38.9%	73.8%
Serious infections	RR = 1.03 (0.26 to 4.06), NS	2.3%	2.4%	7.9%
Infusion reaction	RR = 1.68 (0.56 to 5.04), NS	5.0%	5.4%	NA

NA, not applicable; NS, not significant.

Effectiveness of the technologies compared with newly initiated and previously untried conventional disease-modifying antirheumatic drug

No study addressing the comparison was found.

Effectiveness of the technologies compared with other biologic agents

No study addressing this comparison was found.

Comparison of effectiveness between technologies (head-to-head comparisons)

Evidence from comparative studies

Overview of evidence

One prospective cohort study was identified to compare RTX with TNF inhibitors as a class.[136,137]

Included patients had tried at least one TNF inhibitor (ADA, ETN or IFX) before and discontinued treatment owing to inadequate response. The study was conducted in Switzerland and the median duration of follow-up was 11 months. Full details of this study are provided in *Table 41*.

Patient characteristics

Full details baseline characteristics are reported in *Table 42*. The study included 318 patients and:

- The proportion of women was 77.5%.
- The mean age was 55 years.
- The mean disease duration was 11.3 years.
- The proportion of RF-positive patients was 82.4%.
- Concomitant DMARDs used were MTX (63.9%), LEF (18%) and other (4.5%).
- The proportion of patients receiving steroids was 56.5%.

TABLE 41 Comparative study: characteristics of the included study

Study	Country	Design	Reason for switching	Prior TNF inhibitors; *n*	Treatment arms (no. of patients)	Duration of follow-up	Comments
RCTs							
None were identified							
Non-randomised comparative studies							
Finckh 2009[136,137]	Switzerland	Prospective cohort	Inadequate response	Any (≥ 1)	TNF (163); RTX (155)	11 months (median)	Based on the Swiss Clinical Quality Management program for Rheumatoid Arthritis (SCQOM-RA)
Uncontrolled studies							
Not applicable							

DOI: 10.3310/hta15140

TABLE 42 Comparative study: baseline patient characteristics

Study	Number of patients/% female	Age (years), mean (SD)	RA duration (years), mean (SD)	RF positive (%)	% of patients on concomitant DMARDs and steroids	Number of previous DMARDs, mean	Number of previous TNF inhibitors, mean (SD)	HAQ, mean (SD)	DAS28, mean (SD)	TJC/SJC, mean (SD)	ESR (mm/hour), mean (SD)	CRP (mg/dl), mean (SD)
Finckh 2009[136,137]	318/77.5	55 (12.85)	11.3 (8.1)	82.4	MTX (63.9), LEF (18.0), other (4.5), steroids (56.5)	NR	1 to >2	1.5 (2.9)	4.5 (1.3)	NR	NR	NR

NR, not reported; SD, standard deviation; SJC, swollen joint count; TJC, tender joint count.

- The number of previous DMARDs was not reported.
- The number of previous TNF inhibitors ranged from one to over two.
- The mean baseline HAQ score was 1.5.
- The mean baseline DAS28 score was 4.5.
- No information was provided on CRP and ESR.

Quality assessment

Full details of quality assessment are reported in *Table 43*. The study was a prospective cohort. It had clearly defined inclusion criteria. It was; however, unclear if consecutive patients were included in the study and what percentage of patients were withdrawn.

Results

Table 44 indicates which of the outcomes reported in the main text of the report were assessed in the Finckh *et al*. study.[136,137] No outcomes apart from the ones reported in *Table 44* were assessed.

Withdrawals

Withdrawals were not assessed in this study.

ACR20/50/70 response

ACR response was not assessed in this study.

DAS28

There was a trend favouring TNF inhibitors over RTX for change from baseline in DAS28; however, this difference was not statistically significant (mean difference = –0.35, 95% CI –0.71 to 0.01). The follow-up for this outcome was unclear. See *Figure 83* for details.

EULAR response

EULAR response was not reported in this study.

Health Assessment Questionnaire

Health Assessment Questionnaire score was reported only for baseline in this study.

Quality of life

Quality of life was not reported in this study.

TABLE 43 Comparative study: non-RCT quality assessment

Study	Study design	Inclusion criteria clearly defined?	Were consecutive patients included in the study?	Patients withdrawn (%)	Comments
Finckh 2009[136,137]	Prospective cohort	Yes	Unclear	Unclear	

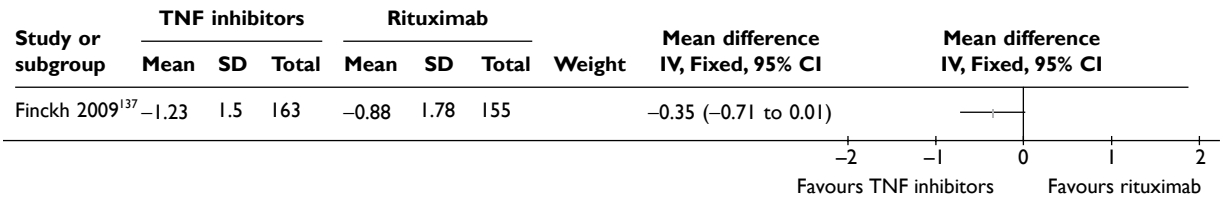


FIGURE 83 Tumour necrosis factor inhibitors versus rituximab: DAS28 change from baseline. SD, standard deviation.

DOI: 10.3310/hta15140

TABLE 44 Comparative study: outcomes assessed in the study and reported in the main text of the report

	Total withdrawal	Withdrawal by reason	ACR (20/50/70)	DAS28	EULAR response	HAQ	QoL	Joint damage	Serious AEs	Infection/ serious infection	Injection/ infusion reaction
Finckh 2009[136,137]				✓							✓

Joint damage

Joint damage was not reported in this study.

Serious adverse events

Serious AEs were not reported in this study.

Infections/serious infections

Infections were not reported in this study.

Injection/infusion reaction

Data for injection/infusion reactions were reported only for a subgroup of 116 patients.[136] Dermatological complications (mainly injection site reactions) occurred in one RTX patient and nine TNF inhibitor patients. Infusion reactions were reported in three RTX and none of the TNF inhibitor patients. Data from both categories were analysed together to compare AEs associated with drug administration (*Figure 84*). There was no statistically significant difference between groups (RR = 1.70, 95% CI 0.56 to 5.22).

Summary

One prospective cohort study[136,137] compared TNF inhibitors as a class with RTX. The median follow-up was 11 months; however, it was not clearly stated when outcomes were assessed. The main results of the study are summarised in *Table 45*.

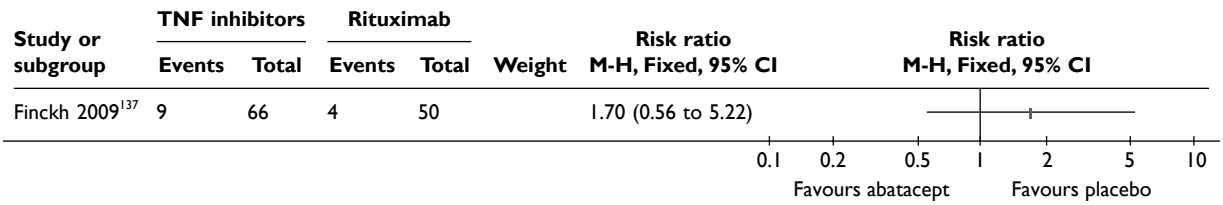


FIGURE 84 Tumour necrosis factor inhibitors versus rituximab: injection/infusion site reactions.

TABLE 45 Comparative study: summary of main results

Outcome	Results (TNF inhibitors vs RTX) Unclear follow-up
Withdrawals	NR
ACR20 response	NR
ACR50 response	NR
ACR70 response	NR
DAS28 – change from baseline	Mean difference = –0.35, 95% CI –0.71 to 0.01, NS
EULAR response	NR
HAQ	NR
QoL	NR
Joint damage	NR
Serious AEs	NR
Any infections	NR
Serious infections	NR
Injection/infusion reactions	RR = 1.70, 95% CI 0.56 to 5.22, NS

NR, not reported; NS, not significant.

Indirect comparisons

Two placebo-controlled RCTs were identified that were considered amenable for an IC of effectiveness of two of the drugs of interest. These trials were REFLEX[124–126] and ATTAIN[127–132] which investigated RTX and ABT, respectively, in similar populations with similar follow-up and outcome measures.

Indirect comparison was conducted (RTX vs ABT) using the method of Bucher *et al.*[73] The following binary outcomes were examined: ACR20, ACR50 and ACR70 responses and 'withdrawal for any reason'. The results are summarised in *Table 46*.

No IC approached statistical significance; however, the IC point estimates slightly favoured RTX for ACR20, ACR70 and withdrawal for any reason.

Indirect comparison for change in HAQ score from baseline to 6 months of treatment was of potential interest. However, data reporting was incomplete in REFLEX[124–126] and the uncertainty in the reported estimates could not be computed reliably. The change in HAQ score was almost the same in the two trials (see *Table 47*) so that it is unlikely that an IC would indicate a difference between the treatments for this outcome measure.

TABLE 46 Indirect comparison: ACR response

Comparison	RR	LCI	UCI	Comment
ACR20				
RTX vs placebo	2.85	2.08	3.91	Favours RTX
ABT vs placebo	2.55	1.74	3.76	Favours ABT
RTX vs ABT	1.12	0.68	1.84	Favours RTX, wide CIs
ACR50				
RTX vs placebo	5.40	2.87	10.16	Favours RTX
ABT vs placebo	5.40	2.21	13.20	Favours ABT
RTX vs ABT	1.00	0.33	2.98	No difference
ACR70				
RTX vs placebo	12.14	2.96	49.86	Favours RTX
ABT vs placebo	6.75	1.63	28.02	Favours ABT
RTX vs ABT	1.80	0.24	13.35	Favours RTX, wide CIs
Withdrawal any reason				
RTX vs placebo	0.39	0.29	0.52	Favours RTX
ABT vs placebo	0.53	0.35	0.81	Favours ABT
RTX vs ABT	0.73	0.44	1.21	Favours RTX, wide CIs

LCI, lower confidence interval; UCI, upper confidence interval.

TABLE 47 Indirect comparison: change from baseline in HAQ score

Change from baseline					
	Active intervention		Placebo		
Study	Mean	SD	Mean	SD	*p*-value
REFLEX[124–126] (RTX)	–0.45	NR	–0.11	NR	$p<0.0001$
ATTAIN[127–132] (ABA)	–0.4	0.6	–0.1	0.5	$p<0.0001$

NR, not reported; SD, standard deviation.

Subgroup analyses

This section summarises results from subgroup analyses. Data from RCTs and observational studies were reported separately. Planned subgroup analyses from placebo-controlled RCTs provide the least biased information with regard to whether *effectiveness* (i.e. the effects of treatment over and above what could be expected without the treatment) varies significantly between the subgroups of interest. Subgroup analyses performed post hoc were highlighted and need to be interpreted with caution.

Owing to the relatively small number of data from RCTs, results from non-randomised, uncontrolled studies were also included but were reported separately from RCT data. Because of the lack of control groups in these studies, any observed differences in the observed *response* (i.e. not corrected for what would happen without treatment) between the subgroups can be due to differences in baseline characteristics before switching, (and the natural course of the disease that follows) as well as genuine differences in the effectiveness between the subgroups.

In accordance with the study selection criteria for non-randomised studies, subgroup analyses were included only if the number of patients was greater than or equal to 20 in at least one of the subgroups being compared. For studies in which some patients were excluded owing to missing data, 'non-responder imputations' were performed and presented for binary outcomes assuming patients with missing data did not achieve the favourable outcomes such as ACR20. 'Observed data' analyses based on actually observed/reported data were presented only when the statistical significance of the results and/or the direction of effect differ from non-responder imputation analyses. For continuous outcomes, results were presented as reported in the original papers and no imputation of missing data was carried out. Where data were available from more than one study for a given outcome/time point, pooled estimates using the random-effects model were presented. Given the potential differences in the populations and methods between studies, the main aim is to illustrate the existence or absence of heterogeneity between studies using the I^2 statistic.

Reasons for withdrawal of the previous tumour necrosis factor inhibitor

Lack of response (primary failure) versus loss of response (secondary failure)

Randomised controlled trials

No evidence from RCTs was reported.

Non-randomised controlled trials

Subgroup data were available for switching to ADA, ETN, an unspecified TNF inhibitor and ABT. No subgroup data were identified for switching to IFX and RTX.

Adalimumab Two uncontrolled studies reported data separately for patients who switched because of lack of response and those who had initial treatment response but later switched because of loss of response.[96,97] Results comparing these two subgroups of patients are summarised in *Tables 48* and *49*.

Overall there was no significant difference in treatment withdrawal between the two subgroups. Patients who switched to ADA because of loss of response had significantly higher response rates for ACR20 and ACR50.

Etanercept Two uncontrolled studies reported subgroup data.[101,104] The results are summarised in *Tables 50* and *51*. Overall the results were similar between the subgroups and no significant difference was observed.

TABLE 48 Switching to ADA owing to lack of response versus owing to loss of response in observational studies – binary outcomes

	Switched due to lack of response		Switched due to loss of response			
Study	*n*/*N*	%	*n*/*N*	%	RR[a] (95% CI)	RD (95% CI)
Withdrawal for any reasons at 3 months						
Bombardieri 2007 (ReAct study)[96]	14/173	8	24/306	8	1.03 (0.55 to 1.94)	0.00 (–0.05 to 0.05)
Withdrawal due to lack of efficacy at 3 months						
Bombardieri 2007 (ReAct study)[96]	5/173	3	5/306	2	1.77 (0.52 to 6.02)	0.01 (–0.02 to 0.04)
Withdrawal due to intolerance/AE at 3 months						
Bombardieri 2007 (ReAct study)[96]	5/173	3	16/306	5	0.55 (0.21 to 1.48)	–0.02 (–0.06 to 0.01)
ACR20 at 3 months						
Bombardieri 2007 (ReAct study)[96]	91/173	53	205/306	67	**0.79 (0.67 to 0.92)**	**–0.14 (–0.24 to –0.05)**
van der Bijl 2008[97]	4/15	27	13/21	62	0.43 (0.17 to 1.06)	**–0.35 (–0.66 to –0.05)**
Pooled estimates					0.69 (0.42 to 1.12)	**–0.20 (–0.37 to –0.02)**
(random effects)					$I^2 = 40\%$	$I^2 = 39\%$
ACR50 at 3 months						
Bombardieri 2007 (ReAct study)[99]	44/173	25	111/306	36	**0.70 (0.52 to 0.94)**	**–0.11 (–0.19 to –0.02)**
van der Bijl 2008[97]	2/15	13	8/21	38	0.35 (0.09 to 1.42)	–0.25 (–0.52 to 0.02)
Pooled estimates (random effects)					**0.68 (0.51 to 0.91)**	**–0.12 (–0.20 to –0.04)**
					$I^2 = 0\%$	$I^2 = 0\%$
ACR70 at 3 months						
Bombardieri 2007 (ReAct study)[96]	15/173	9	41/306	13	0.65 (0.37 to 1.13)	–0.05 (–0.10 to 0.01)
van der Bijl 2008[97]	1/15	7	4/21	19	0.35 (0.04 to 2.83)	–0.12 (–0.33 to 0.09)
Pooled estimates (random effects)					0.62 (0.36 to 1.07)	–0.05 (–0.11 to 0.00)
					$I^2 = 0\%$	$I^2 = 0\%$
EULAR moderate/good response						
Bombardieri 2007 (ReAct study)[96]	127/173	73	243/306	79	0.92 (0.83 to 1.03)	–0.06 (–0.14 to 0.02)
van der Bijl 2008[97]	7/15	47	14/21	67	0.70 (0.38 to 1.30)	–0.20 (–0.52 to 0.12)
Pooled estimates (random effects)					0.92 (0.83 to 1.02)	–0.07 (–0.15 to 0.01)
					$I^2 = 0\%$	$I^2 = 0\%$
EULAR good response						
Bombardieri 2007 (ReAct study)[96]	33/173	19	68/306	22	0.86 (0.59 to 1.24)	–0.03 (–0.11 to 0.04)
van der Bijl 2008[97]	1/15	7	5/21	24	0.28 (0.04 to 2.16)	–0.17 (–0.39 to 0.05)
Pooled estimates (random effects)					0.78 (0.42 to 1.44)	–0.06 (–0.17 to 0.05)
					$I^2 = 11\%$	$I^2 = 28\%$

RD, risk difference; ReAct, Research in Active Rheumatoid Arthritis.

a RR > 1 and RD > 0 favour switch because of loss of response for outcomes related to treatment withdrawal. RR < 1 and RD < 0 favour switch because of loss of response for ACR and EULAR responses.

Bold type indicates statistically significant differences between subgroups.

TABLE 49 Switching to ADA owing to lack of response versus owing to loss of response in observational studies – continuous outcomes

Study	Switch owing to lack of response			Switch owing to loss of response			Mean difference[a] (95% CI)
	N	Mean	SD	*N*	Mean	SD	
DAS28 change from baseline at 3 months							
Bombardieri 2007 (ReAct study)[96]	173	–1.87	1.48	306	–2.03	1.36	0.16 (–0.11 to 0.43)
van der Bijl 2008[97]	15	–1.0	0.9	21	–1.8	2.0	0.80 (–0.17 to 1.77)
Pooled estimates (random effects)							0.30 (–0.22 to 0.83) $I^2 = 36\%$
HAQ change from baseline at 3 months							
Bombardieri 2007 (ReAct study)[96]	173	–0.44	0.54	306	–0.51	0.62	0.07 (–0.04 to 0.18)
van der Bijl 2008[97]	15	–0.13	0.53	21	–0.36	0.48	0.23 (–0.11 to 0.57)
Pooled estimates (random effects)							0.08 (–0.02 to 0.19) $I^2 = 0\%$

ReAct, Research in Active Rheumatoid Arthritis; SD, standard deviation.
a Mean difference > 0 favours switching because of loss of response for DAS28 and HAQ.

TABLE 50 Switching to ETN owing to lack of response versus owing to loss of response in observational studies – binary outcomes

Study	Switched owing to lack of response		Switch owing to loss of response		RR[a] (95% CI)	RD[a] (95% CI)
	n/*N*	%	*n*/*N*	%		
Total withdrawal at 3 months						
Bingham 2009[104]	1/29	3	12/172	7	0.49 (0.07 to 3.66)	–0.04 (–0.11 to 0.04)
ACR20 at 3 months – non-responder imputation						
Buch 2007[101]	14/34	41	13/38	34	1.20 (0.66 to 2.19)	0.07 (–0.15 to 0.29)
Bingham 2009[104]	12/29	41	73/172	42	0.97 (0.61 to 1.56)	–0.01 (–0.20 to 0.18)
Pooled estimates (random effects)					1.06 (0.73 to 1.53) $I^2 = 0\%$	0.02 (–0.12 to 0.17) $I^2 = 0\%$
ACR20 at 3 months – observed data						
Buch 2007[101]	14/34	41	13/38	34	1.20 (0.66 to 2.19)	0.07 (–0.15 to 0.29)
Bingham 2009[104]	12/28	43	73/160	46	0.94 (0.59 to 1.49)	–0.03 (–0.23 to 0.17)
Pooled estimates (random effects)					1.03 (0.72 to 1.48) $I^2 = 0\%$	0.02 (–0.13 to 0.16) $I^2 = 0\%$

TABLE 50 Switching to ETN owing to lack of response versus owing to loss of response in observational studies – binary outcomes *(continued)*

Study	Switched owing to lack of response		Switch owing to loss of response		RR[a] (95% CI)	RD[a] (95% CI)
	n/N	%	*n/N*	%		
ACR50 at 3–4 months – non-responder imputation						
Buch 2007[101]	10/34	29	8/38	21	1.40 (0.62 to 3.13)	0.08 (–0.12 to 0.28)
Bingham 2009[104]	4/29	14	33/172	19	0.72 (0.28 to 1.88)	–0.05 (–0.19 to 0.08)
Pooled estimates (random effects)					1.06 (0.55 to 2.02)	–0.00 (–0.14 to 0.13)
					I^2 = 9%	I^2 = 21%
ACR50 at 3 months – observed data						
Buch 2007[101]	10/34	29	8/38	21	1.40 (0.62 to 3.13)	0.08 (–0.12 to 0.28)
Bingham 2009[104]	4/28	14	33/160	21	0.69 (0.27 to 1.80)	–0.06 (–0.21 to 0.08)
Pooled estimates (random effects)					1.03 (0.52 to 2.05)	–0.01(–0.15 to 0.14)
					I^2 = 19%	I^2 = 29%
ACR70 at 3 months – non-responder imputation						
Buch 2007[101]	5/34	15	5/38	13	1.12 (0.35 to 3.53)	0.02 (–0.14 to 0.18)
Bingham 2009[104]	1/29	3	15/172	9	0.40 (0.05 to 2.88)	–0.05 (–0.13 to 0.03)
Pooled estimate (random effects)					0.86 (0.32 to 2.33)	–0.04 (–0.11 to 0.03)
					I^2 = 0%	I^2 = 0%
ACR70 at 3 months – observed data						
Buch 2007[101]	5/34	15	5/38	13	1.12 (0.35 to 3.53)	0.02 (–0.14 to 0.18)
Bingham 2009[104]	1/28	4	15/160	9	0.38 (0.05 to 2.77)	–0.06 (–0.14 to 0.02)
Pooled estimate (random effects)					0.85 (0.32 to 2.31)	–0.04 (–0.12 to 0.03)
					I^2 = 0%	I^2 = 0%
EULAR good/moderate response at 3 months – non-responder imputation						
Buch 2007[101]	23/34	68	21/38	55	1.22 (0.85 to 1.77)	0.12 (–0.10 to 0.35)
Bingham 2009[104]	17/29	59	100/172	58	1.01 (0.72 to 1.40)	0.00 (–0.19 to 0.20)
Pooled estimates (random effects)					1.10 (0.86 to 1.41)	0.06 (–0.09 to 0.20)
					I^2 = 0%	I^2 = 0%
EULAR good/moderate response at 3 months – observed data						
Buch 2007[101]	23/34	68	21/38	55	1.22 (0.85 to 1.77)	0.12 (–0.10 to 0.35)
Bingham 2009[104]	17/28	61	100/160	63	0.97 (0.70 to 1.34)	–0.02 (–0.21 to 0.18)
Pooled estimate (random effects)					1.07 (0.84 to 1.37)	0.04 (–0.10 to 0.19)
					I^2 = 0%	I^2 = 0%

continued

TABLE 50 Switching to ETN owing to lack of response versus owing to loss of response in observational studies – binary outcomes *(continued)*

Study	Switched owing to lack of response		Switch owing to loss of response		RR[a] (95% CI)	RD[a] (95% CI)
	n/*N*	%	*n*/*N*	%		
EULAR good response at 3 months						
Buch 2007[101]	4/34	12	5/38	13	0.89 (0.26 to 3.06)	–0.01 (–0.17 to 0.14)
Serious AEs						
Bingham 2009[104]	0/29	0	10/172	6	0.27 0.02 to 4.56)	–0.06 (–0.12 to 0.00)
Serious infection						
Bingham 2009[104]	0/29	0	2/172	1	1.15 (0.06 to 23.43)	–0.01 (–0.06 to 0.04)

RD, risk difference.
a RR > 1 and RD > 0 favour switch because of loss of response for outcomes related to treatment withdrawal and AEs. RR < 1 and RD < 0 favours switch because of loss of response for ACR and EULAR responses.

TABLE 51 Switching to ETN owing to lack of response versus owing to loss of response in observational studies – continuous outcomes

Study	Switch due to lack of response			Switch due to loss of response			Mean difference[a] (95% CI)
	N	Mean	SD	*N*	Mean	SD	
DAS28 change from baseline at 3 months							
Buch 2007[101]	34	–1.49	2.25	38	–1.53	2.16	0.04 (–0.98 to 1.06)

SD, standard deviation.
a Mean difference > 0 favour switching because of loss of response for DAS28.

Infliximab No studies of switching to IFX provided subgroup data.

Tumour necrosis factor inhibitors as a class One observational study reported data separately for patients who switched because of lack of response and those who had initial treatment response but later switched because of loss of response.[113] Outcomes for the second TNF inhibitor were reported as an aggregated group and were not reported separately for individual TNF inhibitors. The results from the study are shown in *Tables 52* and *53*.

There were no significant differences between the subgroups in withdrawal and treatment response, except for the analysis with non-responder imputation for good/moderate EULAR response at 3 months. A significantly higher proportion of patients who switched owing to lack of response achieved a good/moderate EULAR response compared with those who switched owing to loss of response. Data were missing for nearly half of the patients in the 'switching owing to loss of response' for several outcomes, which may compromise the reliability of the results.

Rituximab No studies of switching to RTX provided subgroup data.

TABLE 52 Switching to TNF inhibitors owing to lack of response versus owing to loss of response in observational studies – binary outcomes

	Switched owing to lack of response		Switch owing to loss of response			
Study: Blom 2009[113]	*n*/*N*	%	*n*/*N*	%	RR[a] (95% CI)	RD[a] (95% CI)
Withdrawal for any reasons at 3 and 6 months						
3 months	2/49	4	5/75	7	0.61 (0.12 to 3.03)	–0.03 (–0.10 to 0.05)
6 months	6/49	12	16/75	21	0.57 (0.24 to 1.37)	–0.09 (–0.22 to 0.04)
Withdrawal due to lack of efficacy at 3 and 6 months						
3 months	0/49	0	2/75	3	0.30 (0.01 to 6.20)	–0.03 (–0.08 to 0.02)
6 months	4/49	8	10/75	13	0.61 (0.20 to 1.84)	–0.05 (–0.16 to 0.06)
Withdrawal due to intolerance/AE at 3 and 6 months						
3 months	2/49	4	3/75	4	1.02 (0.18 to 5.89)	0.00 (–0.07 to 0.07)
6 months	2/49	4	6/75	8	0.51 (0.11 to 2.43)	–0.04 (–0.12 to 0.04)
EULAR moderate/good response at 3 and 6 months						
3 months – non-responder imputation	25/49	51	16/75	21	**2.39 (1.43 to 4.00)**	**0.30 (0.13 to 0.46)**
3 months – observed data	25/44	57	16/38	42	1.35 (0.86 to 2.12)	0.15 (–0.07 to 0.36)
6 months – non-responder imputation	22/49	45	21/75	28	1.60 (0.99 to 2.58)	0.17 (0.00 to 0.34)
EULAR good response at 3 and 6 months						
3 months – non-responder imputation	7/49	14	3/75	4	3.57 (0.97 to 13.15)	0.10 (0.00 to 0.21)
6 months – non-responder imputation	4/49	8	7/75	9	0.87 (0.27 to 2.83)	–0.01 (–0.11 to 0.09)
DAS28 ≤ 3.2 at 3 and 6 months						
3 months – non-responder imputation	8/49	16	7/75	9	1.75 (0.68 to 4.52)	0.07 (–0.05 to 0.19)
6 months – non-responder imputation	5/49	10	11/75	15	0.70 (0.26 to 1.88)	–0.04 (–0.16 to 0.07)

RD, risk difference.
a RR > 1 and RD > 0 favour switch because of loss of response for outcomes related to treatment withdrawal. RR < 1 and RD < 0 favour switch because of loss of response for EULAR and DAS28-based responses.
Bold type indicates statistically significant differences between subgroups.

TABLE 53 Switching to TNF inhibitors owing to lack of response versus owing to loss of response in observational studies – continuous outcomes

	Switch due to lack of response			Switch due to loss of response			
Study: Blom 2009[113]	N	Mean	SD	N	Mean	SD	Mean difference[a] (95% CI)
DAS28 change from baseline at 3 and 6 months							
3 months (observed data)	44	–1.2	1.0	38	–0.7	1.3	–0.50 (–1.01 to 0.01)
6 months (observed data)	33	–1.3	1.3	41	–0.6	1.3	**–0.70 (–1.30 to –0.10)**

SD, standard deviation.
a Mean difference > 0 favours switching because of loss of response for DAS28.
Bold type indicates statistically significant differences between subgroups.

TABLE 54 Switching to ABT owing to lack of response versus owing to loss of response in the ATTAIN LTE[119] – binary outcomes

Results[a] at 6 months (unless otherwise stated)	Switched due to lack of response		Switched due to loss of response			
	n/N	%	*n*/N	%	RR[b] (95% CI)	RD[b] (95% CI)
ACR20 (non-responder imputation)	73/130	56	50/84	60	0.94 (0.75 to 1.19)	–0.03 (–0.17 to 0.10)
ACR50 (non-responder imputation)	30/130	23	20/84	24	0.97 (0.59 to 1.59)	–0.01 (–0.12 to 0.11)
ACR70 (non-responder imputation)	13/130	10	12/84	14	0.70 (0.34 to 1.46)	–0.04 (–0.13 to 0.05)
HAQ improvement ≥ 0.3 (non-responder imputation)	77/130	59	60/84	71	0.83 (0.68 to 1.01)	–0.12 (–0.25 to 0.01)
HAQ improvement ≥0.3 (observed data)	77/126	61	60/79	76	**0.80 (0.67 to 0.97)**	**–0.15 (–0.28 to –0.02)**
DAS28 ≤3.2 (non-responder imputation) 3 months	11/130	8	11/84	13	0.65 (0.29 to 1.42)	–0.05 (–0.13 to 0.04)
DAS28 ≤3.2 (non-responder imputation) 6 months	21/130	16	17/84	20	0.80 (0.45 to 1.42)	–0.04 (–0.15 to 0.07)
DAS28 <2.6 (non-responder imputation) 3 months	8/130	6	3/84	4	1.72 (0.47 to 6.31)	0.03 (–0.03 to 0.08)
DAS28 <2.6 (non-responder imputation) 6 months	11/130	8	12/84	14	0.59 (0.27 to 1.28)	–0.06 (–0.15 to 0.03)

RD, risk difference.

a Data were reported in the MS to NICE and were not from the published paper.

b RR >1 and RD >0 favour switch because of loss of response for outcomes related to treatment withdrawal. RR <1 and RD <0 favour switch because of loss of response for EULAR and DAS28-based responses.

Bold type indicates statistically significant differences between subgroups.

Abatacept Subgroup data from the LTE of the ATTAIN trial (ATTAIN LTE[119]) were reported in the MS. As patients had to complete 6 months of treatment in the ATTAIN trial[127–132] in order to enter ATTAIN LTE,[119] the included patients were no longer representative of the randomised cohort. The results are shown in *Table 54*. A significant difference between the subgroups was found only in an observed data analysis of HAQ improvement greater than or equal to 0.3 at 6 months. Significantly more patients who switched owing to loss of response achieved this criterion than those who switched owing to lack of response.

Summary

- No conclusion can be made with regard to whether the effectiveness of the five technologies varies according to lack of response or loss of response to the prior TNF inhibitor because of the lack of RCT evidence.
- Evidence from two uncontrolled studies[96,97] of switching to ADA showed significant differences in favour of patients who switched because of loss of response for ACR20 and ACR50.
- Evidence from two uncontrolled studies[101,104] of switching to ETN indicated that there was no significant difference in treatment withdrawal and response between the subgroups.
- Evidence from a Dutch study (DREAM[113]) of switching to an unspecified alternative TNF inhibitor did not find a significant difference between the subgroups.
- Evidence from the ATTAIN LTE[119] of switching to ABT did not find a significant difference between the subgroups except in an analysis based on observed data in which more patients who switched due to loss of response achieved HAQ improvement greater than or equal to 0.3 at 6 months than due to lack of response.

- No evidence from observational studies was identified for switching to IFX and RTX.
- Discussion: there is lack of RCT evidence. It has been speculated that patients who withdrew from a TNF inhibitor owing to lack of response may not respond as well to another TNF inhibitor as those who withdrew owing to loss of response. This was observed in studies of switching to ADA, but not in studies of switching to ETN or an unspecified alternative TNF inhibitor. Of note, a similar trend (higher response rates for patients who withdrew owing to loss of response) was seen in the ATTAIN LTE[119] for switching to ABT, which is not a TNF inhibitor. These observational studies were insufficiently powered to identify clinically important differences and thus the findings require further confirmation.

Switching due to lack of efficacy (lack or loss of response) versus switching due to intolerance (adverse events)

Randomised controlled trials

RCT evidence was available only for RTX. Data were provided in the MS as commercial-in-confidence information.

Rituximab Commercial-in-confidence information (or data) removed.

FIGURE 85 Commercial-in-confidence information (or data) removed.

FIGURE 86 Commercial-in-confidence information (or data) removed.

Non-randomised controlled trials

Subgroup data were available for switching to ADA, ETN and an alternative, unspecified, TNF inhibitor.

Adalimumab Subgroup data were reported in two uncontrolled studies[96,97] and were summarised in *Tables 55* and *56*. The results, mainly driven by the Research in Active Rheumatoid Arthritis (ReAct) study,[96] showed significant differences for EULAR response and change in DAS28 in favour of patients who switched because of intolerance/AEs.

Etanercept Subgroup data were available from one uncontrolled study.[103] The results are presented in *Table 57*. No significant difference between subgroups was found.

Tumour necrosis factor inhibitors as a class Subgroup data were available from three observational studies.[110,112,113] The results are shown in *Tables 58* and *59*. Patients who withdrew from the previous TNF inhibitors because of intolerance/AEs were more likely to withdraw because of intolerance/AEs again compared with those who withdrew from the previous TNF inhibitors because of lack of efficacy. On the other hand, patients who withdrew from the previous TNF inhibitors because of intolerance/AEs were more likely to achieve various ACR, EULAR and other DAS28-based response criteria.

Summary

- Evidence [commercial-in-confidence information (or data) removed]. No subgroup data from RCT were identified for the other technologies.
- Evidence from observational studies was available for switching to ADA, ETN and an alternative, unspecified, TNF inhibitor. Evidence was not available for switching to IFX and ABT.

TABLE 55 Switching to ADA owing to lack of efficacy versus owing to intolerance/AEs in observational studies – binary outcomes

	Switched owing to lack of efficacy		Switched owing to intolerance/AE			
Study	*n*/*N*	%	*n*/*N*	%	RR[a] (95% CI)	RD (95% CI)
Withdrawal for any reasons at 3 months						
Bombardieri 2007 (ReAct)[96]	38/479	8	18/179	10	0.79 (0.46 to 1.35)	–0.02 (–0.07, 0.03)
Withdrawal due to lack of efficacy at 3 months						
Bombardieri 2007 (ReAct)[96]	10/479	2	3/179	2	1.25 (0.35 to 4.47)	0.00 (–0.02 to 0.03)
Withdrawal due to intolerance/AE at 3 months						
Bombardieri 2007 (ReAct)[96]	21/479	4	12/179	7	0.65 (0.33 to 1.30)	–0.02 (–0.06 to 0.02)
ACR20 at 3 months						
Bombardieri 2007 (ReAct)[96]	296/479	62	120/179	67	0.92 (0.81 to 1.04)	–0.05 (–0.13 to 0.03)
van der Bijl 2008[97]	17/36	47	2/5	40	1.18 (0.38 to 3.65)	0.07 (–0.39 to 0.53)
Pooled estimates (random effects)					0.92 (0.82 to 1.05) $I^2 = 0\%$	–0.05 (–0.13 to 0.03) $I^2 = 0\%$
ACR50 at 3 months						
Bombardieri 2007 (ReAct)[96]	155/479	32	68/179	38	0.85 (0.68 to 1.07)	–0.06 (–0.14 to 0.03)
van der Bijl 2008[97]	10/36	28	1/5	20	1.39 (0.22 to 8.66)	0.08 (–0.30 to 0.46)
Pooled estimates (random effects)					0.86 (0.68 to 1.08) $I^2 = 0\%$	–0.05 (–0.13 to 0.03) $I^2 = 0\%$
ACR70 at 3 months						
Bombardieri 2007 (ReAct)[96]	56/479	12	30/179	17	0.70 (0.46 to 1.05)	–0.05 (–0.11 to 0.01)
van der Bijl 2008[97]	5/36	14	0/5	0	1.78 (0.11 to 28.28)	0.14 (–0.11 to 0.39)
Pooled estimates (random effects)					0.71 (0.47 to 1.07) $I^2 = 0\%$	0.00 (–0.17 to 0.17) $I^2 = 53\%$
EULAR good/moderate response						
Bombardieri 2007 (ReAct)[96]	370/479	77	151/179	84	**0.92 (0.85 to 0.99)**	**–0.07 (–0.14 to –0.01)**
van der Bijl 2008[97]	21/36	58	4/5	80	0.73 (0.43 to 1.22)	–0.22 (–0.60 to 0.17)
Pooled estimates (random effects)					**0.91 (0.84 to 0.99)** $I^2 = 0\%$	**–0.08 (–0.14 to –0.01)** $I^2 = 0\%$
EULAR good response						
Bombardieri 2007 (ReAct)[96]	101/479	21	51/179	28	**0.74 (0.55 to 0.99)**	–0.07 (–0.15 to 0.00)
van der Bijl 2008[97]	6/36	17	1/5	20	0.83 (0.12 to 5.57)	–0.03 (–0.40 to 0.34)
Pooled estimates (random effects)					**0.74 (0.56 to 0.99)** $I^2 = 0\%$	–0.07 (–0.15 to 0.00) $I^2 = 0\%$

RD, risk difference.

a RR > 1 and RD > 0 favour switch because of intolerance/AE for outcomes related to treatment withdrawal and AEs. RR < 1 and RD < 0 favour switch because of intolerance/AE for ACR and EULAR responses.

Bold type indicates statistically significant differences between subgroups.

TABLE 56 Switching to ADA owing to lack of efficacy versus owing to intolerance/AEs in observational studies – continuous outcomes

	Switch due to lack of efficacy			Switch due to intolerance/AE			
Study	*N*	Mean	SD	*N*	Mean	SD	Mean difference[a] (95% CI)
DAS28 change from baseline at 3 months							
Bombardieri 2007 (ReAct)[96]	479	–1.97	1.40	179	–2.22	1.28	**0.25 (0.02 to 0.48)**
van der Bijl 2008[97]	36	–1.47	1.64	5	–1.40	0.60	–0.07 (–0.82 to 0.68)
Pooled estimate (random effects)							**0.22 (0.01 to 0.44)**
							$I^2 = 0\%$
HAQ change from baseline at 3 months							
Bombardieri 2007 (ReAct)[96]	479	–0.49	0.59	179	–0.55	0.64	0.06 (–0.05 to 0.17)
van der Bijl 2008[97]	36	–0.26	0.50	5	–0.15	0.34	–0.11 (–0.45 to 0.23)
Pooled estimate (random effects)							0.04 (–0.06 to 0.15)
							$I^2 = 0\%$

SD, standard deviation.
a Mean difference > 0 favours switching because of intolerance/AE for changes in DAS28 and HAQ.
Bold type indicates statistically significant differences between subgroups.

TABLE 57 Switching to ETN owing to lack of efficacy versus owing to intolerance/AEs in an observational study – continuous outcome

	Switch due to lack of efficacy			Switch due to intolerance/AE			
Study	*N*	Mean	SD	*N*	Mean	SD	Mean difference[a] (95% CI)
DAS28 change from baseline (time not specified; between 3 months to 9 months/last observed value on treatment)							
Laas 2008[103]	20	–1.19	2.09	6	–1.30	1.25	0.11 (–1.25 to 1.47)

SD, standard deviation.
a Mean difference > 0 favours switching because of intolerance/AE for changes in DAS28 and HAQ.

TABLE 58 Switching to an alternative TNF inhibitor owing to lack of efficacy versus owing to intolerance/AEs in observational studies – binary outcomes

	Switched owing to lack of efficacy		Switched owing to intolerance/AE			
Study	*n*/*N*	%	*n*/*N*	%	RR[a] (95% CI)	RD[a] (95% CI)
Withdrawal for any reason at 3 and 6 months (non-responder imputation)						
Blom 2009[113] – 3 months	7/124	6	8/73	11	0.52 (0.19 to 1.36)	–0.05 (–0.14 to 0.03)
Blom 2009[113] – 6 months	22/124	18	17/73	23	0.76 (0.43 to 1.34)	–0.06 (–0.17 to 0.06)
Withdrawal due to lack of efficacy at 3 and 6 months (non-responder imputation)						
Blom 2009[113] – 3 months	2/124	2	1/73	1	1.18 (0.11 to 12.76)	0.00 (–0.03 to 0.04)
Blom 2009[113] – 6 months	14/124	11	4/73	5	2.06 (0.70 to 6.02)	0.06 (–0.02 to 0.13)
Withdrawal due to intolerance/AE at 3 and 6 months (non-responder imputation)						
Blom 2009[113] – 3 months	5/124	4	7/73	10	0.42 (0.14 to 1.28)	–0.06 (–0.13 to 0.02)
Blom 2009[113] – 6 months	8/124	6	12/73	16	**0.39 (0.17 to 0.92)**	–0.10 (–0.20 to 0.00)
ACR20 at 3 months (non-responder imputation)						
Karlsson 2008[112]	61/137	45	78/138	57	0.79 (0.62 to 1.00)	–0.12 (–0.24 to 0.00)

continued

TABLE 58 Switching to an alternative TNF inhibitor owing to lack of efficacy versus owing to intolerance/AEs in observational studies – binary outcomes *(continued)*

	Switched owing to lack of efficacy		Switched owing to intolerance/AE			
Study	*n*/*N*	%	*n*/*N*	%	RR[a] (95% CI)	RD[a] (95% CI)
ACR50 at 3 months (non-responder imputation)						
Karlsson 2008[112]	28/137	20	44/138	32	**0.64 (0.43 to 0.97)**	**–0.11 (–0.22 to –0.01)**
ACR70 at 3 months (non-responder imputation)						
Karlsson 2008[112]	8/137	6	10/138	7	0.81 (0.33 to 1.98)	–0.01 (–0.07 to 0.04)
EULAR good/moderate response at 3 months (non-responder imputation)						
Hjardem 2007[110]	38/109	35	19/72	26	1.32 (0.83 to 2.10)	0.08 (–0.05 to 0.22)
Karlsson 2008[112]	80/137	58	100/138	72	**0.81 (0.68 to 0.96)**	**–0.14 (–0.25 to –0.03)**
Blom 2009[113]	41/124	33	21/73	29	1.15 (0.74 to 1.78)	0.04 (–0.09 to 0.18)
Pooled estimate (random effects)					1.02 (0.72 to 1.45) $I^2 = 67\%$	–0.01 (–0.15 to 0.13) $I^2 = 74\%$
EULAR good/moderate response at 6 months (non-responder imputation)						
Blom 2009[113]	43/124	35	21/73	29	1.21 (0.78 to 1.86)	0.06 (–0.07 to 0.19)
EULAR good response at 3 months (non-responder imputation)						
Hjardem 2007[110]	14/109	13	5/72	7	1.85 (0.70 to 4.91)	0.06 (–0.03 to 0.14)
Karlsson 2008[112]	24/137	18	42/138	30	**0.58 (0.37 to 0.90)**	**–0.13 (–0.23 to –0.03)**
Blom 2009[113]	10/124	8	7/73	10	0.84 (0.33 to 2.11)	–0.02 (–0.10 to 0.07)
Pooled estimate (random effects)					0.87 (0.44 to 1.70) $I^2 = 58\%$	–0.03 (–0.13 to 0.08) $I^2 = 77\%$
EULAR good response at 6 months (non-responder imputation)						
Blom 2009[113]	11/124	9	7/73	10	0.93 (0.38 to 2.28)	–0.01 (–0.09 to 0.08)
DAS28 ≤ 3.2 at 3 months (non-responder imputation)						
Karlsson 2008[112]	33/137	24	51/138	37	**0.65 (0.45 to 0.94)**	**–0.13 (–0.24 to –0.02)**
Blom 2009[113]	15/124	12	13/73	18	0.68 (0.34 to 1.35)	–0.06 (–0.16 to 0.05)
Pooled estimate (random effects)					**0.66 (0.48 to 0.91)** $I^2 = 0\%$	**–0.09 (–0.17 to –0.02)** $I^2 = 0\%$
DAS28 ≤ 3.2 at 6 months (non-responder imputation)						
Blom 2009[113]	16/124	13	11/73	15	0.86 (0.42 to 1.74)	–0.02 (–0.12 to 0.08)
DAS28 < 2.6 at 3 months (non-responder imputation)						
Karlsson 2008[112]	16/137	12	25/138	18	0.64 (0.36 to 1.15)	–0.06 (–0.15 to 0.02)

RD, risk difference.

a RR > 1 and RD > 0 favour switch because of intolerance/AE for outcomes related to treatment withdrawal and AEs. RR < 1 and RD < 0 favour switch because of intolerance/AE for ACR and EULAR responses.

Bold type indicates statistically significant differences between subgroups.

- Evidence from two observational studies of switching to ADA showed significant differences for EULAR response and change in DAS28 in favour of patients who switched because of intolerance/AEs.
- No significant difference between subgroups was found in a small, uncontrolled study of switching to ETN.
- Evidence from three observational studies[110,112,113] of switching to an unspecified, alternative TNF inhibitor suggested that patients who withdrew from the previous TNF inhibitor

TABLE 59 Switching to an alternative TNF inhibitor owing to lack of efficacy versus owing to intolerance/AE in observational study – continuous outcome

	Switch due to lack of efficacy			Switch due to intolerance/AE			
Study: Blom 2009[113]	N	Mean	SD	N	Mean	SD	Mean difference[a] (95% CI)
DAS28 change from baseline at 3 and 6 months							
3 months	82	–0.97	1.15	46	–0.80	1.40	–0.17 (–0.65 to 0.31)
6 months	74	–0.91	1.30	40	–1.00	1.40	0.09 (–0.44 to 0.62)

SD, standard deviation.
a Mean difference > 0 favours switching because of intolerance/AE for changes in DAS28 and HAQ.

because of intolerance/AE were more likely to withdraw because of intolerance/AEs and more likely to achieve ACR, EULAR and DAS28-related response criteria than patients who withdrew from the previous TNF inhibitor because of lack of efficacy.

- Discussion: it is suggested that the effectiveness of a TNF inhibitor may differ between patients who have withdrawn from the previous TNF inhibitor because of lack of efficacy and those who have withdrawn because of AEs, but the effectiveness of other technologies with different mechanism of action may not. There is a lack of RCT evidence to confirm the former. RCT evidence suggests that [commercial-in-confidence information (or data) removed]. RCT evidence for ABT is also lacking. Data from observational studies appear to agree with what is expected in terms of treatment withdrawal and treatment response.

Autoantibody status

Randomised controlled trial

RCT data for subgroups stratified by autoantibody status were available only from the REFLEX trial[124–126] of RTX.

Rituximab

Subgroup data stratified by RF status from the REFLEX trial[124–126] were reported in the MS. Randomisation in this trial was stratified by RF status (RF+, defined as a value of RF greater than or equal to 20 IU/ml at screening; or RF–, defined as RF less than 20 IU/ml at screening) and region (US or non-US). The results for ACR20 at 6 months are shown in *Figure 87* (RR) and *Figure 88* [risk difference (RD)] and for all the ACR response criteria are shown in *Table 60*. Although the proportion of patients achieving ACR criteria was generally lower in RF– patients than in RF+ patients, there was no significant difference in treatment effect between the subgroups.

Further subgroup data stratified by baseline RF and anti-CCP status from the REFLEX trial[124–126] were also reported in the MS and are summarised in *Table 61*. Although test for interaction was significant for RD in ACR50, suggesting a greater treatment effect in patients who were either RF or anti-CCP positive than in those with both RF and anti-CCP negative, the number of patients in the latter subgroup was too small to allow firm conclusion to be drawn. This subgroup analysis was performed post hoc and needs to be interpreted with caution.

Non-randomised controlled trials

No subgroup data from observational studies was identified.

Summary

- Evidence from the REFLEX trial[124–126] did not suggest a significant difference in the effectiveness of RTX according to the presence or absence of RF, although the trial may be

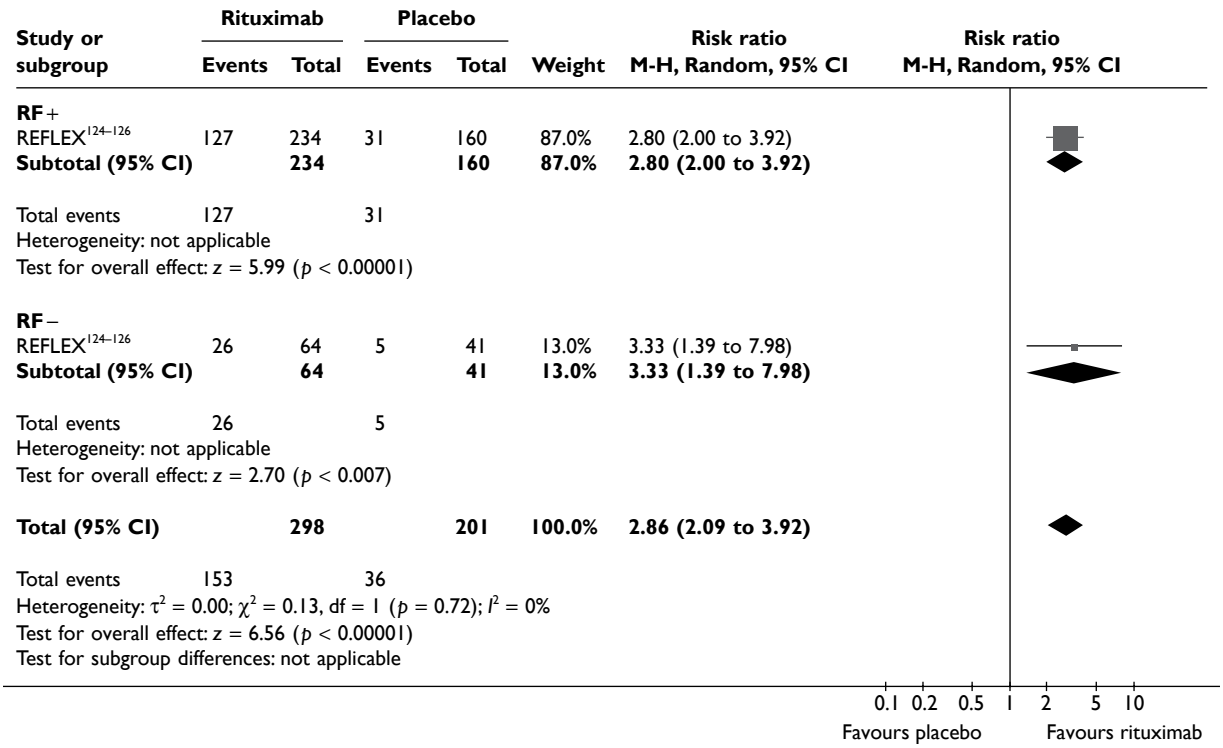


FIGURE 87 Subgroup analysis (switching to rituximab) by RF status: ACR20 at 6 months (relative risk).

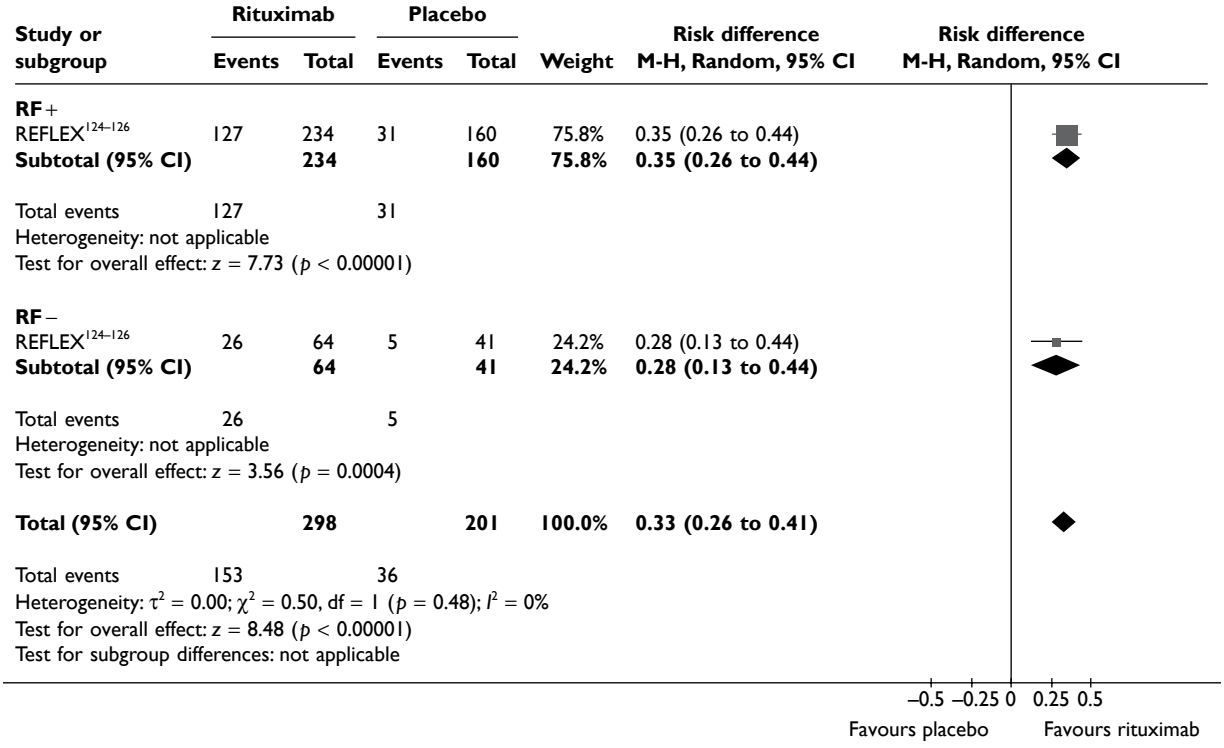


FIGURE 88 Subgroup analysis (switching to rituximab) by RF status: ACR20 at 6 months (risk difference).

TABLE 60 Subgroup analyses (switching to RTX) by RF status in the REFLEX trial:[124–126] ACR20, ACR50 and ACR70 at 6 months

	RTX		Placebo			
Study: REFLEX[124–126]	***n/N***	**%**	***n/N***	**%**	**RR[a] (95% CI)**	**RD (95% CI)**
ACR20 at 6 months						
RF+	127/234	54	31/160	19	**2.80 (2.00 to 3.92)**	**0.35 (0.26 to 0.44)**
RF–	26/64	41	5/41	12	**3.33 (1.39 to 7.98)**	**0.28 (0.13 to 0.44)**
Test for interaction					$p=0.72$	$p=0.48$
ACR50 at 6 months						
RF+	69/234	29	9/160	6	**5.24 (2.70 to 10.19)**	**0.24 (0.17 to 0.31)**
RF–	11/64	17	2/41	5	3.52 (0.82 to 15.09)	**0.12 (0.01 to 0.24)**
Test for interaction					$p=0.63$	$p=0.08$
ACR70 at 6 months						
RF+	31/234	13	3/160	2	**7.07 (2.20 to 22.72)**	**0.11 (0.07 to 0.16)**
RF–	6/64	9	0/41	0	8.40 (0.49 to 145.24)	**0.09 (0.01 to 0.17)**
Test for interaction					$p=0.91$	$p=0.67$

a RR > 1 and RD > 0 favour RTX. RR < 1 and RD < 0 favour placebo.
Bold type indicates statistically significant difference between RTX and placebo within subgroup.

TABLE 61 Subgroup analyses (switching to RTX) by baseline RF and anti-CCP status in the REFLEX trial: [124–126] ACR20, ACR50 and ACR70 at 6 months

	RTX		Placebo			
Study: REFLEX[124–126]	***n/N***	**%**	***n/N***	**%**	**RR[a] (95% CI)**	**RD[a] (95% CI)**
ACR20 at 6 months						
RF and/or anti-CCP positive	79/157	50	19/107	18	**2.83 (1.83 to 4.38)**	**0.33 (0.22 to 0.43)**
RF/anti-CCP negative	8/29	28	1/16	6	4.41 (0.61 to 32.20)	**0.21 (0.01 to 0.41)**
Test for interaction					$p=0.67$	$p=0.33$
ACR50 at 6 months						
RF and/or anti-CCP positive	46/157	29	8/107	7	**3.92 (1.93 to 7.97)**	**0.22 (0.13 to 0.31)**
RF/anti-CCP negative	2/29	7	1/16	6	1.10 (0.11 to 11.25)	0.01 (–0.14 to 0.16)
Test for interaction					$p=0.31$	$\mathbf{p=0.01}$
ACR70 at 6 months						
RF and/or anti-CCP positive	20/157	13	2/107	2	**6.82 (1.63 to 28.55)**	**0.11 (0.05 to 0.17)**
RF/anti-CCP negative	1/29	3	0/16	0	1.70 (0.07 to 39.47)	0.03 (–0.08 to 0.15)
Test for interaction					$p=0.43$	$p=0.24$

a RR > 1 and RD > 0 favour RTX. RR < 1 and RD < 0 favour placebo.
Bold type indicates statistically significant difference between RTX and placebo within subgroup or (for test for interaction) significant difference in treatment effect between subgroups.

underpowered for ruling out a clinically relevant difference between subgroups. There is lack of evidence for other technologies.

- Discussion: in the REFLEX trial,[124–126] the proportion of patients achieving ACR criteria was generally lower in RF– patients than in RF + patients irrespective of treatment group. The treatment effects in terms of RDs between RTX and placebo group were generally larger in RF + patients than in RF– patients, but this does not hold true when RR is used as the measure of effect. Differences between subgroups were not statistically significant according to test for interaction, but the test may be underpowered due to the sample size. Post hoc analysis according to RF and anti-CCP status needs to be interpreted with caution.

Number of tumour necrosis factor inhibitors previously tried

Randomised controlled trials

Randomised controlled trial data stratified by the number of TNF inhibitors the patients had tried before switching were available from the REFLEX trial[124–126] of RTX and the ATTAIN trial[127–132] of ABT.

Rituximab

Subgroup data from the REFLEX trial[124–126] stratified by the number of prior TNF inhibitors (one prior TNF inhibitor vs two or more prior TNF inhibitors) were reported in the MS and are presented in *Table 62*. The results show that RTX was more effective than placebo in both subgroups and there is no significant difference in treatment effects between the subgroups.

Abatacept

Subgroup data from the ATTAIN trial[127–132] stratified by prior TNF inhibitor (ETN, IFX or both) were reported in the MS. For this subgroup analysis, data from patients who had received either ETN or IFX were combined and then were compared with data from patients who had received both ETN and IFX before switching to ABT. The trial was conducted before ADA became widely available and thus few patients had tried more than two TNF inhibitors.

TABLE 62 Subgroup analyses (switching to RTX) by number of prior TNF inhibitors in the REFLEX trial:[124–126] ACR20, ACR50 and ACR70 at 6 months

	RTX		Placebo			
Study: REFLEX[124–126]	*n/N*	%	*n/N*	%	RR[a] (95% CI)	RD[a] (95% CI)
ACR20 at 6 months						
1 prior TNF inhibitor	104/179	58	25/121	21	**2.81 (1.94 to 4.07)**	**0.37 (0.27 to 0.48)**
≥ 2 prior TNF inhibitors	50/119	42	11/80	14	**3.06 (1.70 to 5.50)**	**0.28 (0.17 to 0.40)**
Test for interaction					$p = 0.81$	$p = 0.24$
ACR50 at 6 months						
1 prior TNF inhibitor	54/179	30	8/121	7	**4.56 (2.25 to 9.24)**	**0.24 (0.16 to 0.32)**
≥ 2 prior TNF inhibitors	26/119	22	2/80	3	**8.74 (2.13 to 35.80)**	**0.19 (0.11 to 0.28)**
Test for interaction					$p = 0.41$	$p = 0.46$
ACR70 at 6 months						
1 prior TNF inhibitor	25/179	14	1/121	1	**16.90 (2.32 to 123.06)**	**0.13 (0.08 to 0.18)**
≥ 2 prior TNF inhibitors	12/119	10	2/80	3	4.03 (0.93 to 17.54)	**0.08 (0.01 to 0.14)**
Test for interaction					$p = 0.23$	$p = 0.19$

a RR > 1 and RD > 0 favour RTX. RR < 1 and RD < 0 favour placebo.
Bold type indicates statistically significant difference between RTX and placebo within subgroup or (for test for interaction) significant difference in treatment effect between subgroups.

The results are shown in *Table 63*. Irrespective of the number of prior TNF inhibitor(s), a higher proportion of patients in the ABT group than in the placebo group achieved ACR20 and a HAQ improvement of greater than or equal to 0.3. The difference was larger and statistically significant in the subgroup of patients who had one prior TNF inhibitor, and was smaller and not statistically significant in the subgroup of patients who had two prior TNF inhibitors. The results of tests for interaction do not suggest differential treatment effects between the subgroups, although the tests may be underpowered as the number of patients in the subgroup of two prior TNF inhibitors is relatively small.

Non-randomised controlled trials

Subgroup data stratified by the number of prior TNF inhibitors (or prior biologics) were available for switching to an unspecified TNF inhibitor and to ABT.

Tumour necrosis factor inhibitors as a class

Subgroup data (one prior TNF inhibitor vs two prior TNF inhibitors) were reported in Karlsson *et al.*[112] and the results are presented in *Table 64*. A higher proportion of patients who previously tried one TNF inhibitor achieved various ACR and EULAR response criteria than those who previously tried two TNF inhibitors, although the differences were not statistically significant except for the difference in achieving good EULAR response (25% vs 8%).

In addition to the above, Duftner *et al.*[111] reported a 12-month discontinuation rate of 53.5%, 66.7% (18/27) and 28.6% for the first, second and third biologics (ADA, ETN, IFX and anakinra) in Austrian RA patients. This study included a mixed patient population of those with RA (63%, 109/173) and other rheumatic diseases (37%). The exact number of patients from whom the above RA-specific discontinuation rates were derived was not clearly stated except for the second biologic.

Abatacept

Subgroup data stratified by the number of prior TNF inhibitors (one, two or three) were reported by Schiff *et al.* (ARRIVE study).[120] The results are presented in *Figures 89* and *90*. The results indicate that the proportion of patients achieving DAS28-related response criteria decreases as the number of prior TNF inhibitor(s) that the patients have tried increases (χ^2 test for linear

TABLE 63 Subgroup analyses (switching to ABT) by number of prior TNF inhibitors in the ATTAIN trial:[127–132] ACR20 and HAQ improvement of ≥0.3 at 6 months

	ABT		Placebo			
Study: ATTAIN[127–132]	*n*/*N*	%	*n*/*N*	%	RR[a] (95% CI)	RD[a] (95% CI)
ACR20 at 6 months						
1 prior TNF inhibitor	108/201	54	22/111	20	**2.71 (1.83 to 4.03)**	**0.34 (0.24 to 0.44)**
2 prior TNF inhibitors	21/55	38	4/22	18	2.10 (0.81 to 5.42)	0.20 (–0.01 to 0.41)
Test for interaction					$p=0.63$	$p=0.23$
HAQ improvement from baseline≥ 0.3 at 6 months						
1 prior TNF inhibitor	102/201	51	26/111	23	**2.17 (1.51 to 3.11)**	**0.27 (0.17 to 0.38)**
2 prior TNF inhibitors	19/55	35	5/22	23	1.52 (0.65 to 3.56)	0.12 (–0.10 to 0.33)
Test for interaction					$p=0.45$	$p=0.20$

a RR > 1 and RD > 0 favour ABT. RR < 1 and RD < 0 favour placebo.
Bold type indicates statistically significant difference between ABT and placebo within subgroup or (for test for interaction) significant difference in treatment effect between subgroups.

TABLE 64 Switching to an alternative TNF inhibitor by number of TNF inhibitors previously tried (observational studies) – binary outcomes

	1 prior TNF inhibitor		2 prior TNF inhibitors			
Study: Karlsson 2008[112]	*n*/*N*	%	*n*/*N*	%	RR[a] (95% CI)	RD[a] (95% CI)
ACR20 at 3 months						
	172/337	51	13/36	36	1.41 (0.90 to 2.21)	0.15 (–0.02 to 0.32)
ACR50 at 3 months						
	91/337	27	7/36	19	1.39 (0.70 to 2.76)	0.08 (–0.06 to 0.21)
ACR70 at 3 months						
	24/337	7	1/36	3	2.56 (0.36 to 18.40)	0.04 (–0.02 to 0.10)
EULAR moderate/good response at 3 months						
	240/337	71	21/36	58	1.22 (0.92 to 1.62)	0.13 (–0.04 to 0.30)
EULAR good response at 3 months						
	84/337	25	3/36	8	2.99 (1.00 to 8.98)	**0.17 (0.06 to 0.27)**

a RR > 1 and RD > 0 favour patients who had one prior TNF inhibitor for ACR and EULAR responses.
Bold type indicates statistically significant difference between TNF inhibitors and placebo within subgroup or (for test for interaction) significant difference in treatment effect between subgroups.

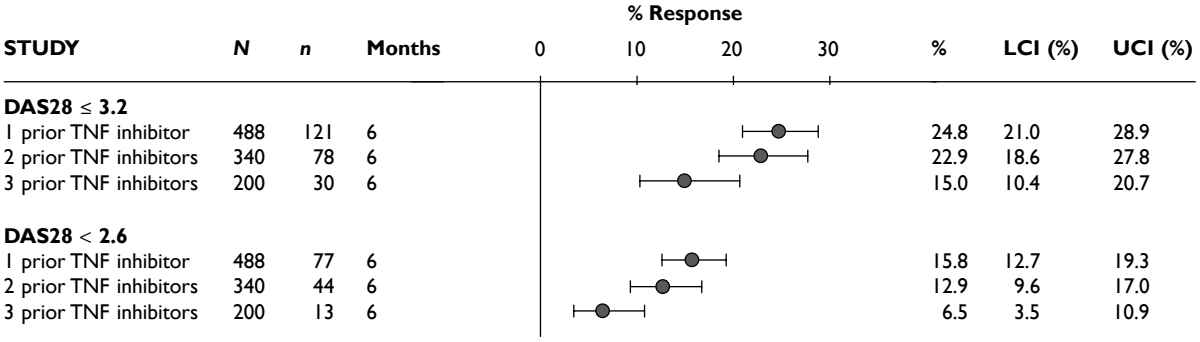


FIGURE 89 Switching to abatacept: DAS28 responses at 6 months stratified by the number of prior TNF inhibitors in the ARRIVE study.[120] LCI, lower confidence interval; UCI, upper confidence interval.

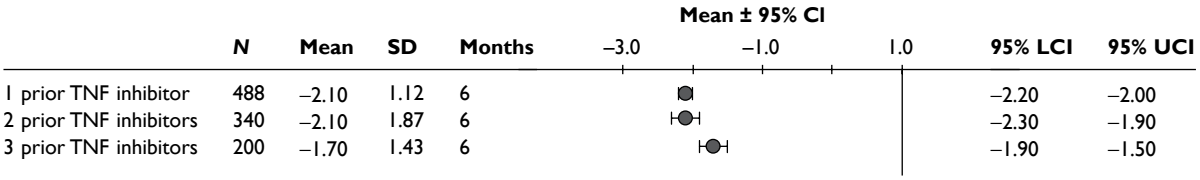


FIGURE 90 Switching to abatacept: DAS28 change from baseline at 6 months stratified by the number of prior TNF inhibitors in the ARRIVE study.[120] LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

trend, $p = 0.009$ for DAS28 less than or equal to 3.2 and $p = 0.005$ for DAS28 less than 2.6). The change in DAS28 from baseline at 6 months was the same for patients who had previously tried one or two TNF inhibitors, but was significantly lower for patients who had previously tried three TNF inhibitors (–2.1 vs –1.7, test for interaction, $p = 0.001$).

Summary

- Evidence from the REFLEX trial[124–126] did not show a significant difference in the effectiveness of RTX (measured as RRs of achieving ACR responses over placebo) between the subgroup of patients who had tried one TNF inhibitor and those who had tried more than one TNF inhibitor. However, the trial may be underpowered for ruling out a clinically relevant difference between the subgroups. The response rates tend to be higher among patients who had tried one TNF inhibitor than among with those who had tried more than one TNF inhibitor irrespective of treatments (i.e. RTX or placebo) received.
- Evidence from the ATTAIN trial[127–132] did not show a significant difference in the effectiveness of ABT (measured as RRs of achieving ACR20 response and HAQ improvement over placebo) between the subgroup of patients who had tried one TNF inhibitor and those who had tried more than one TNF inhibitor. The number of patients in the latter subgroup was small and the difference between ABT and placebo did not reach statistical significance. The trial is likely to be underpowered for ruling out a clinically relevant difference between the subgroups.
- No evidence from RCTs and observational studies was available for the individual TNF inhibitors.
- In an observational study[112] of switching to an unspecified, alternative TNF inhibitor, higher response rates to ACR and EULAR response criteria were reported in patients who tried one TNF inhibitor than in those who tried two TNF inhibitors.
- One observational study[120] of switching to ABT showed that the proportion of patients achieving DAS28-related response criteria decreases as the number of prior TNF inhibitors increases.
- Discussion: many of the studies included in this review included patients who had previously tried more than one TNF inhibitor. Determining whether the effectiveness of the technologies varies depending on the number of TNF inhibitors previously tried is useful to inform the applicability of findings from these studies to the main population of interest for this appraisal, i.e. patients who had previously had inadequate response to one TNF inhibitor. Results from the REFLEX[124–126] and ATTAIN[127–132] trials suggested that the effectiveness of RTX and ABT (measured as RRs of achieving various improvement criteria over placebo) does not differ significantly between patients who have tried one TNF inhibitor compared and those who have tried more than one. The subgroup analyses; however, were limited by the relatively small number of patients, and thus the possibility of differential treatment effect, particularly in terms of RD, cannot be ruled out. Findings from observational studies for switching to an alternative TNF inhibitor and to ABT agree with an inverse relationship between treatment response and number of prior TNF inhibitors. To what extent the effectiveness of the technologies (in particular the TNF inhibitors) varies by the number of prior TNF inhibitors remains unclear owing to the small volume or complete lack of evidence from RCTs.

Prior tumour necrosis factor inhibitor

Randomised controlled trials

RCT data stratified by the TNF inhibitor from which the patients had switched were available only from the ATTAIN trial[127–132] of ABT.

Abatacept

Subgroup data stratified by prior TNF inhibitor (ETN vs IFX) from the ATTAIN trial[127–132] were reported in the MS and are presented in *Table 65*. The results of the subgroup analyses show that ABT is more effective than placebo in both patients who have previously had inadequate response to ETN and those who have previously had inadequate response to IFX. Tests for interaction do not suggest differential treatment effects between subgroups, although the tests may be underpowered.

Non-randomised controlled trials

Adalimumab

Subgroup data stratified by patients who switched from either ETN or IFX to ADA were available from one study (ReAct).[96] The results are shown in *Tables 66* and *67*. No significant difference between the subgroups was found.

In addition to the above, Gomez-Reino *et al.*[108] reported 12-month retention on treatment of 0.75 (95% CI 0.31 to 0.93) for patients who switched from ETN to ADA ($n = 33$) and 0.69 (95% CI 0.43 to 0.85) for patients who switched from IFX to ADA ($n = 14$). No statistical comparisons were made.

Abatacept

Subgroup data stratified by the TNF inhibitor from which the patients switched were reported by Schiff *et al.* (ARRIVE study).[120] The results are presented in *Figures 91* and *92*. At 6 months, there was no significant difference in the proportion of patients who achieved DAS28 less than or equal to 3.2 (χ^2 test, $p = 0.67$) and DAS28 less than 2.6 (χ^2 test, $p = 0.34$). The mean changes from baseline in DAS28 were also similar between the groups (test for interaction, $p = 0.21$).

Summary

- Evidence from the ATTAIN trial[127–132] suggested that the effectiveness of ABT did not vary significantly according to the TNF inhibitor (ETN or IFX) from which the patients had switched, although the subgroup analysis may be underpowered. No RCT evidence was identified for the other technologies.

TABLE 65 Subgroup analyses (switching to ABT) by prior TNF inhibitor (ETN or IFX) in the ATTAIN trial:[127–132] ACR20 and HAQ improvement of ≥0.3 at 6 months

	ABT		Placebo			
Study: ATTAIN[127–132]	***n/N***	**%**	***n/N***	**%**	**RR[a] (95% CI)**	**RD[a] (95% CI)**
ACR20 at 6 months						
Prior ETN	28/61	46	8/43	19	**2.47 (1.25 to 4.88)**	**0.27 (0.10 to 0.44)**
Prior IFX	80/140	57	14/68	21	**2.78 (1.70 to 4.52)**	**0.37 (0.24 to 0.49)**
Test for interaction					$p = 0.78$	$p = 0.39$
HAQ improvement from baseline of ≥ 0.3 at 6 months						
Prior ETN	25/61	41	11/43	26	1.60 (0.89 to 2.90)	0.15 (–0.03 to 0.33)
Prior IFX	77/140	55	15/68	22	**2.49 (1.56 to 3.99)**	**0.33 (0.20 to 0.46)**
Test for interaction					$p = 0.25$	$p = 0.12$

a RR > 1 and RD > 0 favours ABT. RR < 1 and RD < 0 favours placebo.
Bold type indicates statistically significant difference between ABT and placebo within subgroup or (for test for interaction) significant difference in treatment effect between subgroups.

- Evidence from observational studies of switching to ADA[96] and to ABT[120] suggested that treatment response does not vary significantly according to the TNF inhibitor that the patients had previously tried.
- Assuming no interaction between the technologies that have been used sequentially, the results of this subgroup analysis provide an indication of whether patients previously treated with different TNF inhibitors represented distinctly different populations when they switch. Limited data do not suggest this is the case although the evidence is very limited in view of possible combinations of treatment sequence.

TABLE 66 Switching to ADA by prior TNF inhibitor in observational studies – binary outcomes

Study: Bombardieri 2007 (ReAct)[96]	Switched from ETN		Switched from IFX			
	n/N	%	*n/N*	%	RR[a] (95% CI)	RD[a] (95% CI)
Withdrawal for any reasons at 3 months						
	20/188	11	50/591	8	1.26 (0.77 to 2.06)	0.02 (–0.03 to 0.07)
Withdrawal due to lack of efficacy at 3 months						
	5/188	3	12/591	2	1.31 (0.47 to 3.67)	0.01 (–0.02 to 0.03)
Withdrawal due to intolerance/AE at 3 months						
	10/188	5	33/591	6	0.95 (0.48 to 1.90)	0.00 (–0.04 to 0.03)
ACR20 at 3 months						
	107/188	57	378/591	64	0.89 (0.77 to 1.02)	–0.07 (–0.15 to 0.01)
ACR50 at 3 months						
	64/188	34	201/591	34	1.00 (0.80 to 1.26)	0.00 (–0.08 to 0.08)
ACR70 at 3 months						
	24/188	13	77/591	13	0.98 (0.64 to 1.50)	0.00 (–0.06 to 0.05)
EULAR moderate/good response						
	149/188	79	460/591	78	1.02 (0.94 to 1.11)	0.01 (–0.05 to 0.08)
EULAR good response						
	40/188	21	154/591	26	0.82 (0.60 to 1.11)	–0.05 (–0.12 to 0.02)

a RR > 1 and RD > 0 favour switching from IFX for outcomes related to treatment withdrawal. RR < 1 and RD < 0 favour switching from IFX for ACR and EULAR responses.

TABLE 67 Switching to ADA by prior TNF inhibitor in observational studies – continuous outcomes

Study: Bombardieri 2007 (ReAct)[96]	Switched from ETN			Switched from IFX			
	N	Mean	SD	*N*	Mean	SD	Mean difference[a] (95% CI)
DAS28 change from baseline at 3 months							
	188	–2.0	1.4	591	–2.0	1.4	0.00 (–0.23 to 0.23)
HAQ change from baseline at 3 months							
	188	–0.43	0.61	591	–0.51	0.60	0.08 (–0.02 to 0.18)

SD, standard deviation.
a Mean difference > 0 favour switching due to loss of response for DAS28 and HAQ.

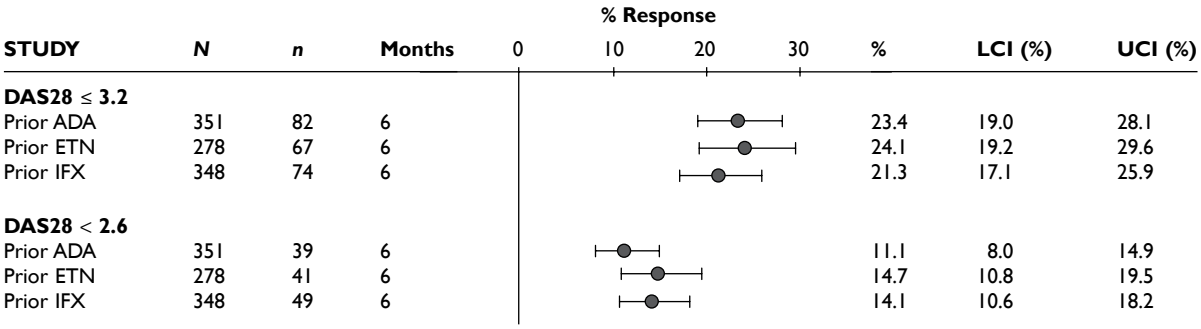


FIGURE 91 Switching to abatacept: DAS28 responses at 6 months stratified by prior TNF inhibitor in the ARRIVE study.[120] LCI, lower confidence interval; UCI, upper confidence interval.

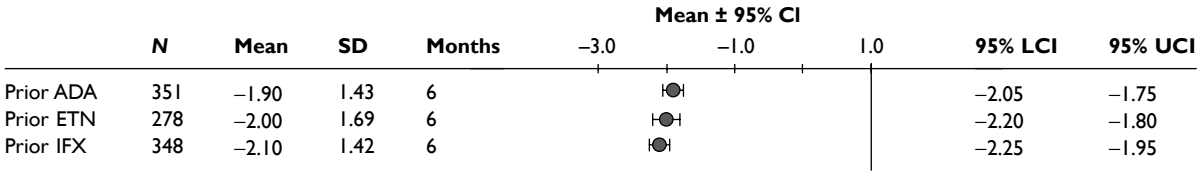


FIGURE 92 Switching to abatacept: DAS28 change from baseline at 6 months stratified by prior TNF inhibitor in the ARRIVE study.[120] LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

Other subgroups

Commercial-in-confidence information (or data) removed.

TABLE 68 Commercial-in-confidence information (or data) removed

Ongoing studies

Electronic searches

Electronic searches for ongoing studies identified only two relevant studies. One of these is looking at extended treatment with RTX in patients who have had an inadequate response (due to toxicity or inadequate efficacy) to previous or current treatment with ETN, IFX or ADA are being entered into an open-label study of two doses RTX and subsequently randomised to a third dose or placebo (if still having B cells). The study acronym is EXTRRA and it is being conducted in the UK. It has a target sample size of 60. The study appears to have been completed in 2010 but has not yet been published. Parts of this study are relevant to the decision problem in this report.

The second study is a 'multicentre clinical observation real-life study' of RTX in patients with active RA whose current treatment with TNF inhibitors in combination with MTX is insufficient. The study acronym is RIRA, and it has a target sample size of 20. It appears to have been undertaken in Austria and to have been completed. This study does not as yet appear to have been published.

Manufacturer's submissions

Mentions of ongoing studies in the MSs were as follows:

- Adalimumab: no explicit statements are provided in the MS about ongoing studies on ADA. Data from large registries are included.

- Etanercept: no explicit statements are provided in the MS about ongoing studies on ETN. Data from registries and LTEs are included.
- Infliximab: the MS provides details on an ongoing multicentre open-label RCT (RE-START; C0168Z05) which aims to assess the efficacy and safety of IFX in patients with active RA who inadequately respond to ETN or ADA. The primary outcome is EULAR response at week 10. Other outcomes will include ACR, tender/swollen joints, HAQ and HRQoL using the SF-36 instrument. Evaluations will be made up to 26 weeks. The study is being conducted in North America, the EU and Israel. The sample size is indicated as ~ 200.
- Rituximab: the MS lists eight ongoing studies (REFLEX open-label extension, SERENE, IMAGE, MIRROR, SUNRISE, SIERRA, DANCER open-label extension, WA16291 and its open-label extension) and various data are presented from these studies in the submission.
- Abatacept: no explicit statements are provided in the MS about ongoing studies on ABT. Data from registries and LTEs are included.

Chapter 4

Assessment of cost-effectiveness

Systematic review of existing cost-effectiveness evidence

Methods

Search strategy

Articles on the cost and cost-effectiveness of drugs for RA after the failure of a TNF inhibitor were identified from the searches for clinical effectiveness. In addition, the NHS Economic Evaluation Database (NHS EED), Cochrane Library 2009 (Issue 3) and the internet sites of national economic units were searched.

Study selection

All articles identified in the searches were imported into the same REFERENCE MANAGER database (REFERENCE MANAGER v.11, Thomson ResearchSoft) as for clinical effectiveness. Titles and abstracts were independently checked for relevance based on the population and intervention by two reviewers alongside selection of papers for clinical effectiveness. If articles were considered relevant by at least one of the reviewers, a full paper copy was ordered. A flow chart presenting the process of selection of studies for the systematic review can be found in *Appendix 3*.

One reviewer applied the inclusion and exclusion criteria using a standard checklist (see *Appendix 7*). Data was extracted by one reviewer using a pre-designed data extraction form and were independently checked by a second reviewer. Data on the following were extracted:

- study characteristics, such as form of economic analysis, population, interventions, comparators, perspective, time horizon and modelling used
- clinical effectiveness and cost parameters, such as effectiveness data, health-state valuations (utilities), resource-use data, unit cost data and key assumptions
- results and sensitivity analyses.

These characteristics and the main results of included economic evaluations are summarised in subsequent tables. The study population and question, selection of alternatives, form of evaluation, effectiveness data, costs, benefit measurement and valuation, decision modelling, discounting, allowance for uncertainty and presentation of results were all evaluated as part of this process.

In addition, all five manufacturers submitted economic analyses. These submissions are reviewed in detail in *Critique of manufacturers' submissions*.

Results

Thirty-eight papers were potentially relevant and ordered. One paper[140] was unobtainable. Four studies met the inclusion criteria and the key features of these studies are summarised in *Table 69*. Further details of the four studies are presented in *Appendix 9*. Their quality was assessed using a simplified version of the Drummond and Jefferson checklist.[141] A summary of the strategies compared and incremental cost-effectiveness ratios (ICERs) reported from these studies is provided in *Table 70*. A list of the excluded papers with reasons for exclusion is presented in *Appendix 5*.

TABLE 69 Summary of published economic analyses

Study	Drug considered	Population (patients with RA who failed to respond adequately to)	Form of economic analysis	Model used	Time horizon
Vera-Llonch 2008[142]	ABT	TNF inhibitors	Cost–utility	Patient-level simulation	10 years Lifetime
Russell 2009[143]	ABT	ETN[a]	Cost-effectiveness	Decision tree	2 years
Kielhorn 2008[144]	RTX	Two non-biologic DMARDs and one TNF inhibitor	Cost–utility	Markov	Lifetime
Lindgren 2009[145]	RTX	One or more TNF inhibitors	Cost–utility	Patient-level simulation	Lifetime

a A strategy of ABT as first biologic was also modelled but this is not relevant to the current review.

TABLE 70 Summary of published ICERs

Drug	Study	Time horizon	Strategies compared	ICER
ABT	Vera-Llonch 2008[142]	10 years	ABT + MTX vs MTX	US$50,576 per QALY
		Lifetime		US$45,979 per QALY
	Russell 2009[143]	2 years	ABT → IFX → DMARDs vs IFX → ADA → DMARDs	CAN$12,514 per additional case of 'low disease-activity state' gained
				CAN$16,829 per additional remission gained
RTX	Kielhorn 2008[144]	Lifetime	RTX → DMARDs vs DMARDs	£14,690 per QALY
			RTX → ADA → IFX → DMARDs vs ADA → IFX → DMARDs	£11,601 per QALY
	Lindgren 2009[145]	Lifetime	RTX → TNF inhibitors vs TNF inhibitors	RTX dominates TNF inhibitors

The review identified two ABT studies, and these differed in how ABT was modelled. Vera-Llonch *et al.*[142] considered ABT with MTX compared with MTX alone while Russell *et al.*[143] considered ABT first, then switch to IFX if there was no response, then switch to conventional DMARDs compared with IFX first, then switch to ADA if there was no response, then switch to conventional DMARDs.

The review also identified two RTX economic evaluations, and these differ in how RTX was modelled. Kielhorn *et al.*[144] considered two different RTX pathways (RTX followed by traditional DMARDs compared with traditional DMARDs only and RTX first, then switch to ADA if there was no response, then switch to IFX if there was no response, then switch to traditional DMARDs, compared with ADA first, then switch to IFX, then switch to conventional DMARDs). Lindgren *et al.*[145] considered RTX first, followed by a series of TNF inhibitors compared with a series of TNF inhibitors.

Data source

Both ABT studies[142,143] used the ATTAIN trial[127–132] as their source for ABT effectiveness. Russell *et al.*[144] also extracted the effectiveness of TNF inhibitors in patients with an inadequate response to TNF inhibitors from the ATTAIN trial,[127–132] assuming a 10% reduction after each switch. The same study also used the TEMPO (Trial of Etanercept and Methotrexate with radiographic Patient Outcomes) trial as the source for ETN effectiveness, when ETN appears in the sequence for the first time in patients with an inadequate response to DMARDs.

The two RTX studies[144,145] used data from the REFLEX trial as their source for RTX effectiveness. Kielhorn *et al.*[144] calculated the mean drop in HAQ score for each of the responder groups from the REFLEX trial.[124–126] Utilities were mapped from the HAQ score and their model uses the equation as estimated by Bansback *et al.*[146] (QoL = 0.76 – 0.28 × HAQ + 0.05 × Female). Lindgren and colleagues[146] in their model mapped utilities from an equation as estimated by patient-level data from the Southern Swedish Arthritis Treatment Group Registry (SSTAG) (QoL = 0.915 – 0.252 × HAQ – 0.05 × Male – 0.107 × DAS28). The SSATG data were also used to estimate the HAQ progression [HAQ progression = 0.106 + 0.241 × (HAQ at treatment start) + 0.002 × (Months on treatment) – 0.087 × (second line) – 0.192 × (third line) – 0.007 × (Disease duration)]. It is unclear though what type of regression was used; the text suggests linear whereas the table suggests logistic.

Study type

Three studies were cost–utility analyses, with the cost-effectiveness ratio reported as cost per QALY gained.[142,144,145] Russell *et al.*[143] used the DAS28 response and reported results in cost per additional case of 'low disease activity state' gained (DAS28 less than 2.6) and cost per additional remission gained (DAS28 less than or equal to 3.2).

Perspective

Kielhorn *et al.*[144] carried out the analysis from the UK health-care perspective. Lindgren *et al.*[145] carried out the analysis from a societal perspective, including direct and indirect costs as well as informal care, therefore, results are not directly relevant to a UK health-care perspective. Vera-Llonch *et al.*[142] carried out the analysis from a third-party payer perspective, including medical treatment only. Finally, Russell *et al.*[143] carried out the analysis from the Swedish health-care perspective. Therefore, results from Russell *et al.*[143] cannot be applied directly to the UK.

Modelling approach

Each study used a different modelling approach. Russell *et al.*[143] used a simple decision-tree structure and modelled cost and outcomes over 2 years. Vera-Llonch *et al.*[142] used a patient simulation model exploring two time horizons: 10 years and lifetime. Kielhorn *et al.*[144] used a Markov model structure with a lifetime time horizon and a 6-month cycle length. Lindgren *et al.*[145] used a patient-level simulation model. The time horizon of the model appears to be lifetime, although this was not explicitly stated in the paper. The model runs for continuous time with no fixed cycle length.

Findings

Russell *et al.*[143] conclude that ABT (followed by IFX, then switch to DMARDs) is a cost-effective strategy in patients with an inadequate response to ETN when compared with IFX (followed by ADA, then switch to DMARDs). The ICER was CAN$12,514 per additional case of 'low disease activity state' gained and CAN$16,829 per additional remission gained. Vera-Llonch *et al.*[142] concluded that ABT (combined with MTX) is cost-effective when compared with MTX alone, with an ICER of US$50,576 per QALY in the 10-year time horizon analysis and an ICER of US$45,979 per QALY in the lifetime time horizon. The results of the ABT studies are not comparable as one study[143] is a cost-effectiveness analysis whereas the other is a cost–utility analysis,[142] the studies do not have the same time horizon and, finally, do not apply the same perspective.

Kielhorn *et al.*[144] concluded that RTX is highly cost-effective for patients who have failed to respond adequately to one biologic DMARD. The ICER for RTX followed by DMARDs was £14,690 per QALY compared with conventional DMARDs only, while the ICER for RTX first, then switch to ADA, then to IFX, then to DMARDs, compared with ADA first, then switch to IFX, then to DMARDs, was £11,601 per QALY. Lindgren *et al.*[145] concluded that the RTX strategy

(followed by a series of TNF inhibitors) was dominant (i.e. cheaper and provided a QALY gain) when compared with a TNF inhibitor strategy. This was explained by the lower price and better effect of RTX than the mix of second-line TNF inhibitors. Both studies favour RTX and their results could be comparable as both studies are cost–utility analyses with a lifetime horizon. However, the study by Lindgren *et al.*[145] uses a societal perspective, which could give a more favourable ICER (in this instance the RTX strategy dominates the TNF inhibitors strategy) as the difference in costs is driven by the indirect costs and the costs of informal care.

Summary

- A direct comparison of ICERs between studies is not possible because of the different approaches to modelling, in particular time horizon, country of origin and perspective chosen.
- All studies used a decision-analytic model. Published models vary in some important aspects: the type of model used, the sequence of drugs, comparator therapies and time horizon.
- Incremental analyses, to which appropriate sensitivity analyses had been applied, were reported without exception.
- All but one study carried out a cost–utility analysis and reported results in 'cost per QALY'. One study carried out a cost-effectiveness analysis and reported results in cost per additional case of 'low disease activity state' gained (DAS28 less than 2.6) and cost per additional remission gained (DAS28 less than or equal to 3.2).
- There was disparity in the selection of perspectives chosen for the analyses. One study reported costs that included both those from a health-care perspective as well as indirect costs and costs of informal care; inclusion of these costs improves the cost-effectiveness of the drug.

Critique of manufacturers' submissions

A submission was received from each company, all including a model-based economic analysis. *Table 71* provides a brief summary of the five economic analyses provided, based on the companies' written submissions.

Abbott submission (adalimumab)

A discrete event simulation model was built to evaluate the cost-effectiveness of ADA. The type of evaluation undertaken was a cost–utility analysis with outcomes measured in QALYs.

Adalimumab was compared with all interventions included in the scope: ETN, IFX, RTX and ABT, all combined with MTX. In each of these five strategies, each drug was followed by GST, then LEF, then CyA, then rescue therapy. A comparison was also made with a strategy of traditional DMARDs only (GST, then LEF, then CyA, then rescue therapy) and also a strategy in which ADA (or ETN) is followed by RTX, GST, then LEF, then CyA, then rescue therapy.

It is assumed that the population has already had an inadequate response to at least two traditional DMARDs, as these are patients who have had an inadequate response to a TNF inhibitor. Therefore, MTX, sulfasalazine and HCQ are not considered as comparators in the economic evaluation.

Response rates are assumed to be equal across all the TNF inhibitors. In addition, drug, administration and monitoring costs of ADA and ETN are assumed to be equal. Therefore, ADA and ETN are evaluated in the same treatment sequence and results for these two drugs are considered similar throughout the submission.

DOI: 10.3310/hta15140

TABLE 71 Summary of methods used in the industry economic analyses

Submission features	ADA (Abbott)	ETN (Wyeth Pharmaceuticals)	IFX (Schering-Plough Ltd)	RTX (Roche)	ABT (Bristol-Myers Squibb Ltd)
Population	Adult patients with active RA who have had an inadequate response to MTX, sulfasalazine, HCQ and one TNF inhibitor	Adult patients with active RA who have had an inadequate response to ETN	Adult patients with active RA who have had an inadequate response to two non-biologic DMARDs and one TNF inhibitor	Adult patients with active RA who have had an inadequate response to a TNF inhibitor	Adult patients with moderate-to-severe RA who have had an inadequate response to at least one TNF inhibitor
Interventions and comparators	**GST** → LEF → CyA → rescue vs **ADA/ETN** → GST → LEF → CyA → rescue vs **IFX** → GST → LEF → CyA → rescue vs **RTX** → GST → LEF → CyA → rescue vs **ABT** → GST → LEF → CyA → rescue vs **ADA/ETN** → **RTX** → GST → LEF → CyA → rescue	**ETN/IFX/ADA** → DMARDs → 'salvage therapy' vs **DMARDs** → DMARDs → 'salvage therapy' vs **RTX** → DMARDs → 'salvage therapy'	**ADA** → DMARDs vs **ETN** → DMARDs vs **IFX** → DMARDs vs **ABT** → DMARDs vs **RTX** → DMARDs vs **ADA** → **RTX** → DMARDs vs **ETN** → **RTX** → DMARDs vs **IFX** → **RTX** → DMARDs vs **DMARDs**	**RTX** → LEF → GST → CyA → palliative care vs **ETN** → LEF → GST → CyA → palliative care vs **ADA** → LEF → GST → c CyA → palliative care vs **IFX** → LEF → GST → CyA → palliative care vs **ABT** → LEF → GST → CyA → palliative care vs **LEF** → GST → CyA → palliative care	**ABT** → **IFX** → LEF → GST → AZA → CyA → penicillamine → palliative care vs **RTX** → **IFX** → LEF → GST → AZA → CyA → penicillamine → palliative care **ABT** → **TNF inhibitors** → LEF → GST → AZA → CyA → penicillamine → palliative care vs **TNF inhibitors** → **TNF inhibitors** → LEF → GST → AZA → CyA → penicillamine → palliative care
Form of analysis	Cost–utility analysis	Cost–utility analysis	Cost–utility analysis	Cost–utility analysis	Cost–utility analysis
Model used	Discrete event simulation model	Markov model	Patient simulation	Patient-level simulation	Patient-level simulation
Cycle length	Continuous	6 months	1 month	6 months	Continuous
Time horizon	Lifetime	Lifetime	Lifetime	Lifetime	Lifetime

continued

TABLE 71 Summary of methods used in the industry economic analyses *(continued)*

Submission features	ADA (Abbott)	ETN (Wyeth Pharmaceuticals)	IFX (Schering-Plough Ltd)	RTX (Roche)	ABT (Bristol-Myers Squibb Ltd)
Base-case results presented: ICERs (£/QALY)	ADA/ETN vs DMARDs: 15,962 IFX vs DMARDs: 21,529 RTX (9 months) vs DMARDs: 10,986 ABT vs DMARDs: 30,104 ADA/ETN → RTX vs DMARDs: 13,797	TNF inhibitors vs DMARDs: 14,501 TNF inhibitors vs RTX: 16,225	ADA vs DMARDs: 35,138 ETN vs DMARDs: 35,898 IFX vs DMARDs: 28,661 ABT vs DMARDs: 44,769 RTX vs DMARDs: 17,422 (9-month dose of RTX); 27,161 (6-month dose of RTX) ADA + RTX vs DMARDs: 27,998 (9-month dose of RTX); 32,345 (6-month dose of RTX) ETN + RTX vs DMARDs: 27,936 (9-month dose of RTX); 32,412 (6-month dose of RTX) IFX + RTX vs DMARDs: 24,236 (9-month dose of RTX); 28,617 (6-month dose of RTX)	RTX vs ETN: RTX dominates RTX vs IFX: RTX dominates RTX vs ABT: RTX dominates RTX vs ADA: 310,771 RTX vs DMARDs: 5,311	ABT → IFX vs RTX → IFX: 20,438 ABT → TNF inhibitors vs TNF inhibitors → TNF inhibitors: 23,019
PSA results	ADA/ETN → RTX vs DMARDs: 100% cost-effective at £30,000/QALY RTX vs DMARDs probability of being cost-effective ~ 60% at £20,000/QALY ADA/ETN vs DMARDs probability of being cost-effective ~ 40% at £20,000/QALY	Not presented	RTX (9-month dose) vs DMARDs: probability of being cost-effective > 90% IFX vs DMARDs: probability of being cost-effective ~ 60% at £30,000/QALY IFX + RTX vs RTX: probability of being cost-effective > 40% at £30,000/QALY	RTX vs ETN: RTX is 100% cost-effective, dominating 74% of iterations RTX vs IFX: RTX is 100% cost-effective, dominating 70% of iterations RTX vs ADA: RTX is 100% cost-effective, dominating 37% of iterations RTX vs ABT: RTX is 100% cost-effective, dominating 70% of iterations RTX vs DMARDs: RTX is 100% cost-effective	Probability of ABT being cost-effective at £30,000/QALY: 99% when compared with RTX 97% when compared with TNF inhibitors

DOI: 10.3310/hta15140

Submission features	ADA (Abbott)	ETN (Wyeth Pharmaceuticals)	IFX (Schering-Plough Ltd)	RTX (Roche)	ABT (Bristol-Myers Squibb Ltd)
HAQ → QoL	EQ-5D = 0.82 − 0.11 × HAQ–0.07 × HAQ^2	EQ-5D = 0.76 − 0.28 × HAQ	NA	EQ-5D = 0.82 − 0.11 × HAQ − 0.07 × HAQ^2	HUI3 = 0.76 − 0.28 × HAQ + 0.05 × female
AEs	Included Rates of TB (for TNF inhibitors) from BSRBR Rates of mild, moderate and serious AEs of ETN, IFX and LEF from an observational study LEF was used as a proxy for all traditional DMARDs ETN was used as a proxy for ADA, ABT and RTX	Included Serious AEs were modelled at £1,181 AEs of conventional DMARDs assumed to be more frequent that those of TNF inhibitors	Although the submission provides background evidence on AEs, they have not been included in the model	Not included The clinical section of the submission indicates that the incidence of AEs is very similar across all treatments in the appraisal	Included Sources were mainly published sources ABT has the lowest rates in all AEs apart from sinusitis
Mortality	Applied a treatment-specific mortality effect. Produced a parametric version of the mortality risk, with adjustments for treatment and HAQ	Used a baseline mortality of 1.63 times general population mortality, with an adjustment for change in HAQ (not clear how they have implemented this as they did not supply their model electronically)	Used mortality ratios dependent on age and gender but no variation by HAQ or treatment	Started from general population mortality and applied a multiplier of 1.33 to the power of the HAQ score, with the parameter 1.33 varied in sensitivity analysis	Started from general population mortality and applied a multiplier of 1.33 to the power of the HAQ score, with the parameter 1.33 varied in sensitivity analysis

EQ-5D, European Quality of Life-5 Dimensions; HUI3, Health Utilities Index Mark 3; PSA, probabilistic sensitivity analysis.

New biologic agents (TOC, golimumab and certolizumab pegol) were excluded from the analysis as these drugs were considered not yet available in the UK.

Adverse events

Adverse events were included in the economic analysis. Rates of TB associated with each of the TNF inhibitors (ADA, ETN, IFX) were based on data from the BSRBR.[147] Rates of mild, moderate and serious AEs were estimated from an observational study in Sweden, which evaluated the safety of patients receiving ETN, IFX or LEF.[148] Values for these drugs were used as proxies for other drugs. The effect of this was that the rate of AEs was higher for conventional DMARDs than for biologics.

Health Assessment Questionnaire to utility

A quadratic mapping mechanism was used in order to convert HAQ scores to European Quality of Life-5 Dimensions (EQ-5D) scores ($EQ\text{-}5D = 0.82 - 0.11 \times HAQ - 0.07 \times HAQ^2$). This equation was estimated through EQ-5D data collected in TOC trials [OPTION (tOcilizumab Pivotal Trial in methotrexate Inadequate respONders) and LITHE (tociLIzumab safety and THE prevention of structural joint damage)].[149] The linear mapping mechanism reported in the same study ($EQ\text{-}5D = 0.89 - 0.28 \times HAQ$) was explored in a sensitivity analysis.

Results

The base-case results show that all drugs (ADA/ETN, IFX, RTX and ABT, all followed by traditional DMARDs) may represent cost-effective treatment options when compared with a sequence of traditional DMARDs. RTX had the lowest ICER (£10,986) while ABT had the highest (£30,104). The strategy of introducing RTX after ADA/ETN (i.e. as a third-line biologic) had an ICER of £13,797 per QALY when compared with traditional DMARDs. The ICERs are as follows:

- ADA/ETN versus DMARDs: £15,962 per QALY
- IFX versus DMARDs: £21,529 per QALY
- RTX (9-month dose) versus DMARDs: £10,986 per QALY
- ABT versus DMARDs: £30,104 per QALY
- ADA/ETN + RTX versus DMARDs: £13,797 per QALY.

Incremental cost-effectiveness ratios of ADA/ETN (followed by DMARDs) versus DMARDs presented in the sensitivity analyses varied from £11,191 per QALY to £26,456 per QALY, with ADA/ETN being cost-effective in the vast majority of the scenarios explored.

The probabilistic sensitivity analysis (PSA) results for 100 replications (for a cohort of 20,000 patients per replication) showed that at a willingness to pay (WTP) of £30,000 per QALY, ADA/ETN followed by RTX is the most cost-effective strategy, with the probability of being cost-effective being close to 1. At a WTP of £20,000 per QALY, RTX followed by conventional DMARDs is cost-effective, with a probability of being cost-effective at around 60%, while there is a 40% (approximate) chance of ADA/ETN followed by conventional DMARDs being cost-effective. The submission; however, states: 'although the cost-effectiveness acceptability curve (CEAC) shows the probability that a treatment sequence is the *most* cost-effective option at various willingness-to-pay thresholds, it does not show all treatment strategies which can be considered cost-effective at these threshold(s)'. Therefore, the submission concludes that although the strategy of ADA/ETN followed by conventional DMARDs is never shown to be cost-effective (submission Figure 3.3.2.1),[150] the deterministic results showed that it is cost-effective, with an

ICER of under £16,000 per QALY. The MS fails to point out though that both RTX followed by conventional DMARDs and ADA/ETN followed by RTX had lower ICERs (£10,986 and £13,797, respectively).

Wyeth Pharmaceuticals submission (etanercept)

A Markov model (6-month cycle) was built to evaluate the cost-effectiveness of ETN. The type of evaluation undertaken was a cost–utility analysis with outcomes measured in QALYs. However, Wyeth Pharmaceuticals did not provide the model that produced the results presented in the submission.

Patients in the model were assumed to receive initial treatment with MTX, then to switch to sulfasalazine, then to switch to a '1st TNF inhibitor'. It is unclear which TNF inhibitor this was. However, cost data suggest that it is ETN in all strategies compared. Therefore, it is assumed that the population modelled were patients whose first failed TNF inhibitor was ETN.

The three strategies compared are second TNF inhibitor, DMARDs and 'rituximab', all followed by traditional DMARDs, and then the 'best supportive care' (salvage therapy). It is unclear; however, which TNF inhibitor is compared in the 'second TNF inhibitor' strategy. Cost data suggest that it was an average of ETN, ADA and IFX combined with MTX. Similarly, in the 'DMARDs' strategy, it was unclear which DMARD was compared: cost data suggest that it was MTX. Finally, the DMARD following a TNF inhibitor seems to be sulfasalazine (again based on cost data).

Cost-effectiveness results were presented for a range of assumed HAQ changes of both the TNF inhibitor (ETN/IFX/ADA) and the conventional DMARDs.

Adverse events

Adverse events were included in the economic analysis. For simplicity, only serious AEs were modelled, assuming that they last for one cycle (6 months) only. The cost of a serious AE was estimated at £1,181, which included two general practitioner (GP) visits and 7 inpatient days. The text (submission p. 33) suggests that various published sources were used for the rates of AEs for each drug. AE rates for all TNF inhibitors were assumed to be the same for ETN. Data in the table suggest that rates of AEs are higher in traditional DMARDs than in biologics.

Health Assessment Questionnaire to utility

A linear mapping mechanism was used in order to convert HAQ scores to EQ-5D scores during each model cycle (EQ-5D = 0.76 – 0.28 × HAQ).[151] It was assumed that patients experiencing serious AEs would lose 0.05 units of utility (or 10% of a QALY) over 1 year.

Results

Results were presented for a range of assumed HAQ changes of both TNF inhibitor (ETN/IFX/ADA) and conventional DMARDs. The ICER for TNF inhibitors versus conventional DMARDs was £14,501, when a HAQ drop of 0.55 was assumed for the TNF inhibitors and no change was assumed for the conventional DMARDs. The ICER for TNF inhibitors versus RTX was £16,225, when a HAQ drop of 0.55 was assumed for the TNF inhibitors and a HAQ drop of 0.40 was assumed for RTX.

Probabilistic sensitivity analysis results were not presented in the submission.

Schering-Plough Ltd submission (infliximab)

A patient-simulation/individual sampling model was used to evaluate the cost-effectiveness of IFX. The type of evaluation undertaken was a cost–utility analysis with outcomes measured in QALYs.

Nine treatment sequences were compared in the cost-effectiveness analysis:

- ADA/ETN/IFX/RTX/ABT, each followed by a sequence of traditional DMARDs
- ADA/ETN/IFX, each followed by RTX and then a sequence of traditional DMARDs
- a sequence of traditional DMARDs.

Patients in the model could receive a maximum of two biologic DMARDs followed by a maximum of three non-biologic DMARDs and were limited to a maximum of five treatments within each of the nine sequences. New biologic agents (TOC, golimumab and certolizumab pegol) are excluded from the analysis as these drugs were considered not yet available in the UK.

The baseline characteristics of patients in the GO-AFTER (GOlimulab After Former anti-tumour necrosis factor Therapy Evaluated in Rheumatoid arthritis) trial, in which treatment with a TNF inhibitor (golimumab) following withdrawal from one or more previous TNF inhibitors (ADA, ETN or IFX) was investigated, were considered for the start of the model.

Adverse events

Adverse events were not included in the model although evidence on AEs was included in the efficiency part of the submission.

Health Assessment Questionnaire to utility

There was no mapping mechanism applied on EQ-5D scores. Utility gains or losses were modelled directly using a QoL measure. Each treatment was associated with an initial utility gain, which was estimated from BSRBR data.

Results

The base-case results showed that ADA, ETN, IFX and RTX (followed by traditional DMARDs) might represent cost-effective treatment options, whereas ABT (followed by traditional DMARDs) did not represent a cost-effective treatment option, when all strategies are compared with a sequence of traditional DMARDs. The ICERs were as follows:

- ADA versus DMARDs: £35,138 per QALY
- ETN versus DMARDs: £35,898 per QALY
- IFX versus DMARDs: £28,661 per QALY
- ABT versus DMARDs: £44,769 per QALY
- RTX (9-month dose) versus DMARDs: £17,422 per QALY
- RTX (6-month dose) versus DMARDs: £27,161 per QALY.

Further analysis, adding RTX after the TNF inhibitors (ADA, ETN, IFX), was performed. IFX had the lowest ICER for both doses of RTX explored (6-month dose/9-month dose) when compared with both traditional DMARDs and RTX (both followed by traditional DMARDs). The ICERs were as follows:

- Versus DMARDs:
  - ADA + RTX (9-month dose): £27,998 per QALY
  - ADA + RTX (6-month dose): £32,345 per QALY
  - ETN + RTX (9-month dose): £27,936 per QALY

- ETN + RTX (6-month dose): £32,412 per QALY
- IFX + RTX (9-month dose): £24,236 per QALY
- IFX + RTX (6-month dose): £28,617 per QALY.

- Versus RTX:
  - ADA + RTX (9-month dose): £41,747 per QALY
  - ADA + RTX (6-month dose): £39,084 per QALY
  - ETN + RTX (9-month dose): £42,477 per QALY
  - ETN + RTX (6-month dose): £39,673 per QALY
  - IFX + RTX (9-month dose): £33,274 per QALY
  - IFX + RTX (6-month dose): £30,549 per QALY.

Overall, when compared with DMARDs, RTX had the lowest ICER for both 9-month (£17,422 per QALY) and 6-month doses (£27,161 per QALY). Among TNF inhibitors (ETN, IFX, ADA), IFX had the lowest ICER (£28,661 per QALY).

Incremental cost-effectiveness ratios in the sensitivity analyses varied from £16,752 per QALY (RTX vs DMARDS, when a HAQ improvement of 0.01 per annum was assumed for all biologic DMARDS) to £58,850 per QALY (IFX + RTX vs RTX, when the weight of the patient was assumed to be 120 kg).

The PSA results showed that, when compared with traditional DMARDs, the probability of RTX (9-month dose) being cost-effective was greater than 90% at a range of WTP thresholds greater than £20,000 per QALY. When a 6-month dose was assumed for RTX, the probability of RTX being cost-effective was marginally greater than the probability of IFX being cost-effective, at WTP greater than £20,000 per QALY. The probability of IFX (vs DMARDs) being cost-effective was ~ 60% at £30,000 per QALY. When compared with RTX, the probability of IFX followed by RTX being cost-effective was greater than 40% at £30,000 per QALY.

Roche submission (rituximab)

A patient-level simulation was built to evaluate the cost-effectiveness of RTX. The type of evaluation undertaken was a cost–utility analysis with outcomes measured in QALYs.

Rituximab was compared with all interventions included in the scope: ADA, ETN, IFX and ABT. In addition, RTX was compared with a strategy of traditional DMARDs. In all strategies compared, the first active treatment was followed by salvage therapy consisting of LEF, GST and CyA followed by palliative care. Response rates of LEF, GST and CyA were assumed to be equivalent to MTX for this population. Comparison of RTX against the new biological agents (TOC, golimumab and certolizumab pegol) was not performed as these treatments were considered not used in routine clinical practice in the NHS.

Adverse events

Adverse events were not included in the economic analysis. The clinical section of the submission indicates that the incidence of AEs was very similar across all treatments in the appraisal. Given that RTX was compared head-to-head with each of the interventions in the scope, it was assumed that the costs of treating an AE would be the same in all strategies compared and therefore the cost-effectiveness ratios would not be affected by these costs.

Health Assessment Questionnaire to utility

A quadratic mapping mechanism was used in order to convert HAQ scores to EQ-5D scores during each model cycle ($EQ\text{-}5D = 0.82 - 0.11 \times HAQ - 0.07 \times HAQ^2$). This equation was estimated through EQ-5D data collected in two Roche phase III trials [DMARD-inadequate response (IR)]

for TOC. The linear mapping mechanism used by Bansback *et al.*[146] [Health Utilities Index Mark 3 (HUI3) = 0.76 – 0.28 × HAQ + 0.05 × female] was explored in a scenario analysis.

The model also assumed that the relationship of HAQ score to patient-reported utility was independent of the number of previous biologics used. Moreover, for the base-case analysis, the model allowed for estimates of QALYs being less than zero, when patients progress to very high HAQ scores. However, this relationship was not explored in the sensitivity analysis by adding a restriction to the negative QALY values.

Results

The base-case results showed that RTX dominates ETN (incremental costs –£13,246; incremental QALYs 0.0168), IFX (incremental costs –£10,490; incremental QALYs 0.0699) and ABT (incremental costs –£16,075; incremental QALYs 0.0606). When compared with ADA, RTX was less costly (incremental costs –£13,551) but also less effective (incremental QALYs –0.0436) with an ICER of £310,771 per QALY. When compared with the traditional DMARDs strategy, RTX was more costly (incremental costs £6,323) but also more effective (incremental QALYs 1.0705), with an ICER of £5,311 per QALY.

Overall, TNF inhibitors (ETN, IFX, ADA) were dominated by RTX, i.e. RTX was more effective and less costly. ADA was marginally more effective but also more costly than RTX, resulting in an ICER of £310,771 per QALY. When compared with traditional DMARDs, RTX was cost-effective at £5,311 per QALY.

Incremental cost-effectiveness ratios in the sensitivity analyses varied from £4,898 per QALY (vs traditional DMARDs when a 9-month time to retreatment was assumed for RTX) to £326,397 per QALY (vs ADA when a linear mapping mechanism was assumed for the HAQ to QoL conversion), while in most of the scenarios RTX dominated the other strategies (i.e. RTX was less costly and more effective).

The PSA results for 1,000 Monte Carlo simulations showed that the probability of RTX being cost-effective is 100% at a wide range of WTP thresholds (£5,000–400,000 per QALY).

Bristol-Myers Squibb Ltd submission (abatacept)

A patient-level simulation model was built to evaluate the cost-effectiveness of ABT. The type of evaluation undertaken was a cost–utility analysis with outcomes measured in QALYs. Baseline patient characteristics were from the ATTAIN trial.[127–132] Data from ATTAIN,[127–132] REFLEX[124–126] and BSRBR were used for the treatment efficacy of the drugs modelled.

Abatacept was compared with all interventions included in the scope: ADA, ETN, IFX and RTX. However, TNF inhibitors were also grouped under a 'basket' of TNF inhibitors and these were the base-case comparator. The rationale was reported as based on the conclusions from the NICE appraisal of the sequential use of TNF inhibitors.[152] In addition, the submission argued that TNF inhibitors were grouped because no data were available to draw conclusions about the efficacy of different TNF inhibitors, after a failure of a first TNF inhibitor.

The 'basket' labelled TNF inhibitors was defined through use of market share data estimated through survey data (Bristol-Myers Squibb Ltd data on file). These were 22% ETN, 52% ADA, 24% IFX and 2% RTX for the second-line treatment, and 15% ETN, 9% ADA, 37% IFX and 38% RTX for the third-line treatment, as presented on p. 134 of the submission. Patients in the model were randomly assigned to one of the three 'basket' treatments, based on these data, after excluding RTX. Efficacy, costs and other parameters related to that therapy were applied to the

proportion of patients receiving that therapy. Total costs and outcomes of the 'basket' treatment are the sum of the three 'basket' therapies.

There were two main comparisons. In the first comparison ABT was compared with RTX, both followed by IFX, then traditional DMARDs, then palliative care. In the second comparison, ABT was compared with a 'basket' of TNF inhibitors, both followed by another 'basket' of TNF inhibitors, then traditional DMARDs, then palliative care.

Traditional DMARDs were not considered as comparators in the economic analysis on the basis that the target population (RA patients with an inadequate response to TNF inhibitors) should have tried multiple traditional DMARDs, and so it was assumed that clinicians were unlikely to revert to these therapies. DMARDs were included as part of the sequence of treatments only after an insufficient response or intolerance to multiple biological therapies (after failure of three biologic DMARDs). After failing DMARDs, patients received NSAIDs only (palliative care).

Other new biologic agents were not considered as comparators for two reasons. Firstly, price information for the new biological therapies was not available at the time of writing. Secondly, new biological therapies were considered not routinely used in the NHS.

In summary, this submission did not consider a 'non-biologic' strategy. All strategies compared included at least two biologic DMARDs (patients with an inadequate response to one TNF inhibitor).

Adverse events

Adverse events were assumed to reduce QoL as well as increasing costs. The following AEs were included in the economic analysis: infusion-related reaction, injection site reactions, upper respiratory tract infection and urinary tract infection, rash, nausea, neutropenia, hypotension, leucopenia, severe allergic reaction and sinusitis. The sources for the rates of the AEs were mainly published data.[126,130] ABT was associated with the lowest rates of all AEs apart from sinusitis.

Health Assessment Questionnaire to utility

A linear mapping mechanism was used in order to convert HAQ scores to HUI3 scores during each model cycle (HUI3 = 0.76 – 0.28 × HAQ + 0.05 × female).[153] The submission discussed the available sources for conversion of HAQ to utility, and selected the formula above for the base-case analysis, on the basis that this formula was used in previous RA appraisals and models[144,146,154] and was preferred over other algorithms[155,156] by the Evidence Review Group (ERG) in the original ABT appraisal. The submission acknowledged that the average baseline HAQ score of 1.5 from the formula selected might not be appropriate for a population with an inadequate response to one TNF inhibitor, and therefore explored the EQ-5D approach[157] in sensitivity analysis.

Results

The base-case results showed that ABT was cost-effective when compared with RTX (both followed by IFX as the third biologic) with an ICER of £20,438 per QALY. ABT was also cost-effective when compared with a 'basket' of TNF inhibitors (both followed by another 'basket' of TNF inhibitors) with an ICER of £23,019 per QALY. Overall, the results showed that the ICERs for ABT were all below £30,000 whether compared with single or a 'basket' of TNF inhibitors, or RTX.

Incremental cost-effectiveness ratios for ABT in the sensitivity analyses varied from £14,145 per QALY (vs RTX, when a 1.5% discount rate was assumed for QALYs) to £40,534 (vs RTX, when the ABT HAQ progression rate was assumed to be 0.012 rather than –0.013 in the base case).

The PSA results showed that the probability of ABT being cost-effective was 99% at £30,000 per QALY when compared with RTX. When compared with a 'basket' of TNF inhibitors, the probability of ABT being cost-effective was 97% at £30,000 per QALY. However, the submission failed to report any other PSA results (particularly below the £30,000 per QALY threshold). From the presented figures it seems that at £20,000 per QALY, both RTX and the 'basket' of TNF inhibitors were cost-effective when compared with ABT, with the probabilities being greater than 50% and greater than 95%, respectively.

Summary

A key issue is the appropriate comparator to be used. All but one submission chose conventional DMARDs as their base-case comparator. One submission did not consider a strategy of conventional DMARDs at all, assuming a switch to a third biologic in all strategies compared.

All submissions used the same type of economic evaluation, with cost per QALY being offered as efficiency measure.

There is some variation in the methods used and sources of data for important model inputs such as QoL scores or baseline population characteristics. Three submissions considered AEs in their model; however, methods and sources of rates and costs of AEs varied.

Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions

Four of the manufacturers (Abbott, Schering-Plough Ltd, Roche and Bristol-Myers Squibb Ltd) used results from ICs and/or mixed-treatment comparisons (MTCs) to inform their model. This section provides a critical appraisal of these analyses and highlights issues that may impact on the validity of their results.

Before commencing on the critique of IC/MTC, it is pertinent to clarify the definition of these terms. NICE's Methods guide (2008) states that 'a mixed treatment comparison (MTC) includes trials that compare the interventions head-to-head and indirectly', whereas an IC is a 'synthesis of data from a network of trials'. These two terms have been used inconsistently and sometimes inter-changeably in some of the MSs. In this section of the assessment report, all the syntheses of data from a network of trials without incorporating evidence from head-to-head trials are referred to as ICs in line with the methods guidance. This also avoids creating a false impression that direct evidence from head-to-head trials was included in these analyses. Only analyses that incorporated both direct and indirect evidence were referred to as MTCs.

For the RA population defined in the scope of this appraisal (patients who had inadequate response to a TNF inhibitor), no head-to-head trial between the five technologies under assessment was identified by the assessment group and the manufacturers, and thus it was not possible to carry out an MTC. IC was possible between RTX and ABT through placebo-controlled trials of respective drugs. This was conducted by Roche and Bristol-Myers Squibb Ltd.

Owing to the lack of head-to-head trials and a complete absence of placebo-controlled trials for the three TNF inhibitors under assessment in the population defined by the scope, three manufacturers have attempted to carry out ICs/MTCs by extending the inclusion criteria to the RA population outside the scope (e.g. patients who had not been treated with a TNF inhibitor and/or patients who had not been treated with MTX). One head-to-head trial exists in this broader population and thus an MTC combining direct and indirect evidence is possible. The key issue for this approach is whether basic assumptions with regard to clinical and methodological homogeneity and exchangeability of estimated treatment effects between trials held.

Indirect comparisons in patient population specified in the scope

Roche and Bristol-Myers Squibb Ltd performed ICs for the RA population defined in the scope using network meta-analyses/Bayesian methods (see *Table 72*, Roche TNF-IR IC and Bristol-Myers Squibb Ltd IC). The ICs were based on the same placebo-controlled RCTs for RTX (REFLEX trial[124–126]) and ABT (ATTAIN trial[127–132]), and additionally included a placebo-controlled RCT for TOC [RADIATE (Research on Actemra Determining effIcacy after Anti-TNF failurEs)]. A further golimumab RCT (GO-AFTER) was also in the Bristol-Myers Squibb Ltd analysis. No placebo-controlled trial for the patient population defined in the scope was identified for ADA, ETN and IFX, and thus it was not possible to include the three TNF inhibitors. The selection and inclusion of TOC and golimumab trials in the IC seemed arbitrary as they provided no evidence regarding the relative effectiveness of RTX versus ABT. The inclusion of these trials had little impact on the estimates of relative effectiveness (expressed as response rates to ACR response criteria and RRs/odds ratios) between RTX and ABT compared with results from a pair-wise IC conducted by the assessment group based on the same RTX and ABT trials (see bottom of *Table 73*).

Roche and Bristol-Myers Squibb Ltd used results from ICs described above to inform their model (ACR responses for Roche; HAQ changes for Bristol-Myers Squibb Ltd). However, this was restricted to the estimates of effectiveness for RTX and ABT and was not applicable for the estimates of effectiveness for TNF inhibitors. For TNF inhibitors, Roche used results from a separate MTC based on different patient populations outside the scope (described below), whereas Bristol-Myers Squibb Ltd used observational data from the BSRBR. The comparisons of effectiveness between TNF inhibitors and RTX/ABT in their models were therefore *not* based on an IC or MTC.

Mixed-treatment comparisons in patient population outside the scope

Three manufacturers have carried out MTCs based on RCTs of an RA population outside the scope [e.g. patients who had not been treated with a TNF inhibitor and/or patients who had not been treated with MTX; see *Table 72* Abbott (MTC), Schering-Plough Ltd (MTC) and Roche (DMARD-IR MTC)].

Owing to the broad inclusion criteria beyond the scope of the appraisal, substantial clinical and statistical heterogeneity exists between the RCTs included in the MTCs. The basic requirement for ICs/MTCs regarding the exchangeability of relative treatment effects between the included studies could not be assumed and thus the validity of the results was questionable. The violation of the basic requirement was particularly prominent in the MTCs conducted by Abbott and Schering-Plough Ltd, in which RCTs of early RA patients who were naive to MTX treatment were included in the analyses along with RCTs of late RA patients who had inadequate response to MTX and/or TNF inhibitors.

Despite the broad inclusion criteria for the MTCs, clinical and methodological similarity/difference between the included studies was only briefly described or not mentioned at all. Statistical heterogeneity between included studies was either not assessed or (where assessed) only dealt with by using random-effects models without further exploration of the potential source of heterogeneity. All the MTCs included a head-to-head trial [ATTEST (Abatacept or infliximab vs placebo, a Trial for Tolerability, Efficacy and Safety in Treating RA), comparing IFX with ABT], but did not examine the direct evidence separately from indirect evidence. Consistency between direct and indirect evidence was not examined.

There is an appreciable difference between the results obtained from the three MTCs (which were based on population outside the scope) and the actual results (where available) observed in RCTs conducted in relevant populations defined in the scope (see *Table 73*). For ACR response criteria, results from these MTCs tend to overestimate the response rates (for both intervention

TABLE 72 Summary of ICs/MTCs reported in MS

Summary item	Abbott (ADA) MTC	Schering-Plough Ltd (IFX) MTC	Roche (RTX): TNF-IR IC and DMARD-IR MTC	Bristol-Myers Squibb Ltd (ABT) IC
Literature search	Based on a number of previous studies/ reports (Nixon *et al.* 2007,[158] Wailoo 2008,[159] Bristol-Myers Squibb Ltd submission[160] and ERG report for TA141[161]) plus an updated search of PUBMED from 1 January 2005 to 31 May 2009	Search of EMBASE, MEDLINE, MEDLINE In Process and Cochrane Library from inception to April 2009; bibliographies of identified studies	Search of MEDLINE and EMBASE from 1990 through 2007	Search of multiple databases from 1 January 1990 to 8 May 2009, conference abstracts, manufacturers and NICE web sites, bibliographies of identified studies
Inclusion criteria	*Design*: clinical trial *Population*: broader than scope (including patients not previous treated with TNF inhibitors and/or MTX) *Intervention*: broader than scope (including anakinra, certolizumab pegol, golimumab, and TOC) *Outcome*: need to report ACR response *Other*: at least 6-month follow-up time	*Design*: double-blind RCTs ≥ 24 weeks (except RTX trials) *Population*: broader than scope (including RA patients of any stage) *Intervention*: broader than scope (including certolizumab pegol, golimumab and TOC) *Outcome:* need to report ACR response criteria or mean change in HAQ score *Other*: published as full papers in English	*Design*: RCTs of duration ≥ 6 months *Population*: two analyses were performed: ***TNF inadequate response (TNF-IR) IC***: same as scope ***DMARD inadequate response (DMARD-IR) MTC***: *Population*: outside scope (including patients who had inadequate response to DMARD but predominantly not previously treated with a TNF inhibitor) *Intervention*: broader than scope (including TOC) *Outcome*: need to report ACR response criteria/ACR core disease parameters *Other*: published as full papers in English, German, French and Dutch	*Design*: RCTs *Population*: same as scope intervention: broader than scope (including certolizumab pegol, golimumab and TOC) *Outcome*: clinically relevant outcomes *Other*: published as full papers in English and conducted in Europe or America
Included studies	**29 RCTs**, plus one open-label randomised study, three prospective cohort study, one study based on registry **Within scope (2)**: ABT (1), RTX (1), plus five other studies of TNF inhibitors **Outside scope (27)**[c]: ABT (4), ADA (5) ETN (5), IFX (2), anakinra (3), certolizumab pegol (1), golimumab (3), TOC (5)	**34 RCTs** **Within scope (2)**: RTX (1), ABT (1) **Outside scope (32)**[c]: ADA (7), ETN (5), IFX (4), RTX (2), ABT (4), certolizumab pegol (3), golimumab (3), TOC (5)	***TNF-IR IC*: Three RCTs** **Within scope (2)**: RTX (1), ABT (1) **Outside scope (1)**: TOC (1) ***DMARD-IR MTC*: 18 RCTs** **Within scope (0)**: none **Outside scope (18)**:[a,c] ADA (4), ETN (4), IFX (3), RTX (2),[b] ABT (3), TOC (3)	**Four RCTs** **Within scope (2)**: RTX (1), ABT (1) **Outside scope (2)**: TOC (1) and golimumab (1)
Assessment of homogeneity and similarity[d] between included studies	Not stated	Not stated. Plots of the treatment effect on ACR response against baseline HAQ and disease duration were used to selected covariables into the analyses	Homogeneity at each ACR response level was assessed using *Q*-statistics. Stated that 'baseline characteristics across the trials were comparable with respect to ACR core parameters'	Not stated
Outcome analysed	ACR20, ACR50 and ACR70	ACR20, ACR50 and ACR70	ACR20, ACR50 and ACR70	Multiple outcomes including ACR responses; response criteria derived from DAS HAQ scores; withdrawal, DAS and HAQ change from baseline; various outcomes on AEs, component outcomes of ACR criteria; SF-36 component summary scores

TABLE 72 Summary of ICs/MTCs reported in MS *(continued)*

Summary item	Abbott (ADA) MTC	Schering-Plough Ltd (IFX) MTC	Roche (RTX): TNF-IR IC and DMARD-IR MTC	Bristol-Myers Squibb Ltd (ABT) IC
Analytical methods	Bayesian hierarchical models estimated with WINBUGS. ACR responses were modelled on a log-odds ratio scale. Log-odds ratios of responses were adjusted for addition of MTX, disease duration and baseline HAQ among other variables. Also used 'Fully conditional predictive mean matching' to impute data	Network meta-analyses conducted on an ordered logit scale. Analyses were performed both with and without adjustment of disease duration	Analyses were performed with WINBUGS and conducted with non-informative priors. Results for TNF inhibitors were pooled	Models were fitted using WINBUGS, employing Markov chain Monte Carlo (MCMC) simulation. Both fixed-effects and random-effects estimation was conducted for all analyses
Input into the manufacturer model	Using Bayesian hierarchical models, posterior mean predicted treatment response rates (predicted for a patient with a disease duration of 11 years and an average HAQ score of 2.1)	Odds ratios (adjusted for disease duration) for ACR responses derived from IC were used in the model	For RTX and ABT, ACR response rates from TNF-IR ICs were used. For TNF inhibitors, ACR response rates from DMARD-IR MTC were firstly discounted by 30% and then used in the model	Results from IC for HAQ change were used in the model, but only for RTX and ABT. Data from registry (BSRBR) on HAQ change were used for TNF inhibitors
Comments	Included trials of both early and late RA populations with very different treatment history (e.g. patients who had inadequate response to a TNF inhibitor vs patients who were naive to TNF inhibitors vs patients who were naive to MTX). The basic requirement for ICs with regard to exchangeability of relative treatment effect between trials cannot be assumed and thus the validity of the results is questionable Also the IC included evidence from multiple study design (i.e. RCTs and observational studies). RCT evidence did not appear to have been analysed separately from evidence from observational studies. The nature of randomised comparison therefore may not have been preserved. In addition, different search strategies and inclusion criteria were applied for different technologies	Included trials of both early and late RA populations with very different treatment history (e.g. patients who had inadequate response to a TNF inhibitor vs patients who were naive to TNF inhibitors vs patients who were naive to MTX). The basic requirement for ICs with regard to exchangeability of relative treatment effect between trials cannot be assumed. The validity of the results is questionable particularly because the IC used MTX as the reference standard (i.e. the hub of the evidence network) for comparison The proportional odds assumption of the ordered logit model (i.e. treatment effect was constant across ACR20, 50 and 70) did not seem to be consistent with observations from REFLEX[124–126] and ATTAIN[127–132] trials	Patient populations included in TNF-IR IC were in line with the scope. The major limitation of the analysis was that only one trial each was available for RTX and ABT and no trial was available for the three TNF inhibitors The inclusion of the TOC trial appeared arbitrary as it provided no information regarding relative effectiveness of RTX and ABT. The inclusion of the trial had little impact on the estimates of relative effectiveness (in terms of ACR responses) between RTX and ABT compared with a pair-wise adjusted IC conducted by the assessment group based on the same trials (see bottom of *Table 73*) RRs were translated into response rates using the pooled placebo response as baseline. Given the substantial heterogeneity between studies (e.g. placebo response rates for ACR20 ranged from 15% to 72% according to Figure 35 of Roche submission), the validity of pooling placebo response across studies and consequently the RRs derived from it was questionable	Patient populations included in the IC were in line with the scope. The major limitation of the analysis was that only one trial each was available for RTX and ABT and no trial was available for the three TNF inhibitors The inclusion of the TOC and golimumab trials appeared arbitrary as they provided no information regarding relative effectiveness of RTX and ABT. The inclusion of these trials had little impact on the estimates of relative effectiveness (in terms of ACR responses) between RTX and ABT compared with a pair-wise adjusted IC conducted by the assessment group based on the same trials (see bottom of *Table 73*)

TA, technology appraisal.

a Four studies were excluded from main analyses (but included in sensitivity analyses) because the 'treatment arms in these trials were fundamentally different from the remaining trials': no DMARD background treatment was provided in three studies; the other study evaluated combination therapy with a biologic agent and sulfasalazine.

b Approximately one-third of patients in this study had previously been treated with a TNF inhibitor.

c One trial included both ABT and IFX.

d As described in Song *et al.* 2009.[162]

TABLE 73 Comparison of ACR responses between data observed in RCTs and results of ICs and MTCs

	Interventions/comparators[a]	ACR20	ACR50	ACR70
ACR responses				
	Control (traditional DMARD/placebo/none)	%		
Data from RCTs	GO-AFTER (week 14)	18	6	2
	REFLEX (week 24)	18	5	1
	ATTAIN (week 24)	20	4	2
Results from IC/MTC	Abbott MTC (model input)	25	10	4
	Roche DMARD-IR MTC	32	12	4
	Roche TNF-IR IC	15	4	1
	TNF inhibitors	%		
Data from RCT	GO-AFTER (golimumab 50 mg) week 24	34	18	12
	GO-AFTER (golimumab 100 mg) week 24	44	20	10
Results from IC/MTC	Abbott MTC (model input)	64	40	21
	Roche DMARD-IR MTC			
	ADA	66	44	18
	ETN	64	36	14
	IFX	60	33	14
	30% degradation of Roche DMARD-IR MTC (model input)	%		
	ADA	46	31	13
	ETN	45	25	10
	IFX	42	23	10
	RTX	%		
Data from RCT	REFLEX (RTX) week 24	51	27	12
Results from IC/MTC	Abbott MTC (model input)	62	38	20
	Roche DMARD-IR MTC	60	35	18
	Roche TNF-IR IC (model input)	46	23	14
	ABT	%		
Data from RCT	ATTAIN (ABT) week 24	50	20	10
Results from IC/MTC	Abbott MTC (model input)	55	31	15
	Roche DMARD-IR MTC	59	33	15
	Roche TNF-IR IC (model input)	43	22	8
Estimates of relative effectiveness				
	TNF inhibitors vs control (odds ratios)			
Data from RCT	GO-AFTER (golimumab 50 mg) week 24	2.55	4.12	4.0
	GO-AFTER (golimumab 100 mg) week 24	3.87	4.67	3.5
Results from IC/MTC	Schering-Plough Ltd MTC, ADA	Commercial-in-confidence information (or data) removed		
	Schering-Plough Ltd MTC, ETN	Commercial-in-confidence information (or data) removed		
	Schering-Plough Ltd MTC, IFX	Commercial-in-confidence information (or data) removed		
	Bristol-Myers Squibb Ltd IC, golimumab 50 mg	2.55	4.30	NA

TABLE 73 Comparison of ACR responses between data observed in RCTs and results of ICs and MTCs *(continued)*

	Interventions/comparators[a]	ACR20	ACR50	ACR70
	Bristol-Myers Squibb Ltd IC, golimumab 100 mg	3.90	4.89	NA
	RTX vs control (odds ratios)			
Data from RCT	REFLEX (RTX) week 24	4.77	7.00	13.67
Results from IC/MTC	Schering-Plough Ltd MTC	Commercial-in-confidence information (or data) removed		
	Bristol-Myers Squibb Ltd IC	4.84	7.27	16.38
	ABT vs control (odds ratios)			
Data from RCT	ATTAIN (ABT) week 24	4.18	6.53	7.40
Results from IC/MTC	Schering-Plough Ltd MTC	Commercial-in-confidence information (or data) removed		
	Bristol-Myers Squibb Ltd IC	4.20	6.98	9.28
	RTX vs ABT (RRs)			
Results from ICs	Assessment group IC	1.12	1.00	1.80
	Roche TNF-IR IC	1.06	1.05	1.75
	Bristol-Myers Squibb Ltd IC	1.14	1.07	1.85
	(ratio of odds ratios)			

IR, inadequate response; NA, not applicable.
a All interventions and comparators were assumed to be used with ongoing MTX.
Commercial-in-confidence information (or data) removed.

and control arms but to a different extent) compared with the response rates observed in relevant RCTs.

The substantial heterogeneity among studies included in these MTCs and the discrepancy between the results from these analyses and those actually observed in RCTs raise serious concern with regard to the validity of the MTCs as well as the validity of economic evaluations that utilised data from them.

Further critique of manufacturers' models

A description of the models included in each of the MSs and a summary of results from this modelling is provided in section *Critique of manufacturers' submissions*. A critique of ICs and/or MTCs that were used to inform the models is given in section *Critique of manufacturers' submissions*. Building upon *the Critique of manufacturers' submissions*, this section aims to provide further critique of the manufacturers' models by highlighting issues and uncertainties related to data input and assumptions used.

Data input and assumptions used in the manufacturer models are summarised in *Table 74*. Key issues relating to characteristics of starting population, estimates of clinical effectiveness (short term and long term), mapping of effectiveness data to utility, discontinuation rule(s) and treatment duration, handling of AEs and mortality, estimates of costs and other relevant factors are discussed below for each of the models.

Abbott (adalimumab)

Characteristics of starting population

The characteristics of the starting population were based on data from the BSRBR[122] that is appropriate. These published data were collected in 2006 and are slightly dated. The starting population in the Abbott model had a slightly higher HAQ score at baseline than the equivalent population described in the current British Society for Rheumatology (BSR) submission (2.1 vs 2.0). The current BSR submission to NICE (Section 4, Table 4–1)[163] highlights a trend over the past 8 years that patients treated more recently have shorter disease durations, lower DASs, and lower HAQ scores and have tried fewer conventional DMARDs before starting a TNF inhibitor.

Treatment sequence

The stated assumptions that patients will have tried MTX, sulfasalazine and HCQ (and thus these drugs are not evaluated) are clinically appropriate. The evaluated sequences include gold as the comparator or first traditional DMARD after failing biologics (see *Table 71*). Sequences that consider GST early are increasingly unlikely. GST is now likely to be used much later during treatment (for example, Survey of West Midlands rheumatologists, *Appendix 11*). In addition, although AZA has limited efficacy, this drug would still be tried in patients with resistant disease. This drug should therefore be used late in the sequence.

Estimates of clinical effectiveness – short term

Clinical effectiveness was estimated according to ACR response rates obtained from the manufacturer's MTC, which included RCTs of very heterogeneous patient populations outside the scope of this appraisal as well as a few selected observational studies of relevant populations within the scope. As described in section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*, the validity of the MTC was questionable. The ACR responses estimated from the MTC for control groups (i.e. placebo or DMARDs for which patients had had inadequate response) were used for conventional DMARDs in the model. These response rates, if estimated correctly, would not have reflected the response rates for a conventional DMARD that patients had not previously tried.

Mapping of ACR responses to HAQ change was based on an RCT (DE019) of ADA used as the *first* biologic therapy. Mapping using alternative data from PREMIER (an RCT of ADA in early RA, MTX-naive patients) suggested that the relationship between ACR response to treatment and changes in HAQ score will differ depending on the population being treated. Therefore, mapping based on data from a subgroup of patients in DE019 with a HAQ score greater than 2 was used by the manufacturer in a sensitivity analysis.

Estimates of clinical effectiveness – long term

The base case assumed that HAQ progression on biologics is the same as that of the general population (0.03 per year). An annual increase of 0.045 for conventional DMARDs and 0.06 for non-responders was assumed. Zero HAQ progression on biologic treatment was explored in sensitivity analyses. While previous analyses have considered the possibility that HAQ does not progress at all in a population of patients treated with a TNF inhibitor this assumption lacks face validity. Remission was achieved by 7% of patients in a large cohort of RA patients and minimum disease activity was achieved by around 20%, including those on a TNF inhibitor.[164] On the basis that a majority of RA patients treated with a TNF inhibitor have continued disease activity, it is not credible that HAQ does not change with time in this population.

The model assumed that, following treatment withdrawal, the HAQ score would immediately worsen by an exactly equivalent amount to the initial improvement. A sensitivity analysis was conducted in which the HAQ score worsens by 75% of the initial gain. It seems appropriate to explore several possible scenarios. Patients experiencing a severe flare of disease are unlikely to

be left in this state and unlikely to suffer a prolonged worsening of function because of the short-term use of corticosteroids combined with other DMARDs and/or a biologic as appropriate.

Mapping of effectiveness data to utility

Health Assessment Questionnaire scores were converted to EQ-5D scores according to equations developed by Ducournau *et al.*,[159] using data from TOC trials [OPTION (tOcilizumab Pivotal Trial in methotrexate Inadequate respONders) and LITHE (tociLIzumab safety and THE prevention of structural joint damage)] in patients who had had inadequate response to MTX. Two equations (linear and non-linear) were available. The non-linear equation was used for the base-case analysis, while the linear equation was examined in sensitivity analyses.

Discontinuation rule and treatment duration

The model demands for an ACR50 response at 6 months in order that patients are eligible to continue treatment. This threshold appears too high compared with clinical practice. It is clear from the BSRBR and other data that patients continue treatment with a TNF inhibitor despite not meeting NICE-stipulated DAS28 criteria (so called 'stayers' in BSRBR analyses). This suggests that there is worthwhile clinical benefit despite a failure to meet thresholds (which are derived from populations and have limitations when applied to individual patients; see *Chapter 1*, *Disease Activity Score response criteria*).

Withdrawal rates used in the base-case analysis for TNF inhibitors are based on a shared frailty model previously developed by the Decision Support Unit using BSRBR data for patients receiving their second TNF inhibitor. Withdrawal rates for ABT and RTX were assumed to be the same as for TNF inhibitors.

Handling of adverse events and mortality

A reduction in mortality (independent of age, HAQ and comorbidity) for patients on TNF inhibitors was assumed based on Jacobsson *et al.*[166] This assumption was also applied for RTX and ABT. A hazard ratio (HR) of 0.92 for males and 0.52 for females was used. The reported mortality advantages for patients on TNF inhibitor treatment compared with conventional DMARDs need great care in interpretation because of selection biases involved in treating patients with a TNF inhibitor which may not be sufficiently adjusted for. Sicker individuals, those with cardiac failure and those with previous malignancies are much less likely to be treated.

Estimates of costs

Abbott states that the drug costs of ADA and ETN are similar but fails to acknowledge that this applies only to ADA used every other week. The licence for ADA permits dose increases so the drug may be administered every week (potentially doubling drug costs). European data, including from the UK, suggest that around 8% of patients need an increase in their dose of ADA. This figure may be an underestimate as many investigators reported that financial constraints inhibited dose increases.[167]

The dose of LEF is 20 mg per day not 25 mg as stated. The stated dose for CyA was 2.5 mg per kg. In practice, this can range from 2.5 mg to 4 mg per kg.

The stated six outpatient visits and 11 nurse visits during the first 6 months for patients starting a TNF inhibitor are excessive for ETN and ADA. For IFX the necessary assessments can be done on the day a patient receives an infusion though it may be appropriate to include a nurse visit at other times to ensure that MTX safety is maintained. So, there will be five visits for infusions during the first 6 months. Blood and other monitoring can be done at these visits and an additional two nurse visits would be needed to ensure that a minimum of monthly checks were made.

Two outpatient visits and six nurse visits were assumed for monitoring after the first 6 months. An outpatient visit every 3 months is appropriate for a period of around 18 months, but after this, in stable patients with well-controlled disease, monitoring by a rheumatologist can be reduced to every 6 months. Frequency of blood testing for concomitant MTX can be done at nurse visits or in GP practices where there are shared care agreements.

Disease-related hospital costs (inpatient days and joint replacement procedures) were estimated based on HAQ band using data from the Norfolk Arthritis Register (NOAR) database.[168] Higher costs are more likely with higher HAQ scores, but for items such as joint replacement this is likely to apply only to those with persistently raised HAQ scores (i.e. those with more fixed damage) rather than in those whom HAQ scores rise as a result of flares of inflammatory disease. The latter group have a higher risk of hospitalisation because of this but rates in contemporary practice are low because of the use of corticosteroids.

Wyeth Pharmaceuticals (etanercept)

Wyeth Pharmaceuticals did not submit an electronic version of the model. Overall, the description with regard to methods for identifying data and justification for the selection of data was very limited.

Characteristics of starting population

The mean age of the starting population was 53 years and was based on the TEMPO trial. This is an RCT of TNF-naive patients (mean disease duration 6.6 years) who had not experienced treatment failure with MTX. The rationale for choosing this trial is not described. The modelling appears to start when patients first receive RA treatment (MTX), so it is not clear why a starting cohort of early RA patients was not chosen. The starting population in TEMPO was younger than the BSRBR cohort at study entry (mean age 56 years), but it is difficult to ascertain whether patients' age would be similar to the BSRBR data (i.e. reflecting UK population and practice) when the patients reached the point of failing a TNF inhibitor. Other characteristics of the starting population were not described, including baseline HAQ score.

Treatment sequence

The identity of drugs in the treatment sequence was not clearly described. For example, the terms 'first TNF-α inhibitor', 'second TNF-α inhibitor' and 'DMARD after TNF' were used without further clarification. The costs for the second TNF inhibitor (the intervention under evaluation) were assumed to be the average of ADA, ETN and IFX + MTX. The assumed costs for the second TNF inhibitor (£4,159.68), therefore, do not reflect the (higher) costs for ETN + MTX (£4,687.83) according to the table of unit costs provided in the Wyeth Pharmaceuticals submission.

Estimates of clinical effectiveness – short term

Short-term HAQ improvement for TNF inhibitor (average –0.48; varied between –0.55 to –0.41 depending on reasons for withdrawal of previous TNF inhibitor) was based on data from the ReAct study,[96] an observational study of switching to ADA after failing a TNF inhibitor. Short-term HAQ improvement for conventional DMARD was assumed to be zero according to the BSRBR data. In contrast with the –0.48 observed in ReAct study,[96] short-term HAQ improvement for TNF inhibitor observed in BSRBR was only –0.11, but these data were not used in the model. The estimates of effectiveness for the model were therefore taken from studies using different methods of data collection and thus inappropriate for comparison.

Various sources have been cited for HAQ improvements on other treatments but the citations may be incorrect (e.g. the cited references for DMARDs before first TNF inhibitor appears to be uncontrolled studies of second-line biologics).

Estimates of clinical effectiveness – long-term

Long-term HAQ progression for patients on TNF inhibitors (and RTX) was assumed to be zero according to Wick *et al.*[93] Various levels of HAQ progression were applied for patients on conventional DMARDs based on assumption.

Mapping of effectiveness data to utility

The HAQ score was converted to EQ-5D score using the equation reported by Brennan *et al.*[151]

Discontinuation rule and treatment duration

This was not described.

Handling of adverse events and mortality

Various probabilities of experiencing a serious AE were assigned for each treatment. The cited references included a systematic review including probably first-line biologic use, narrative reviews and methodological papers discussing HAQ and QoL (possibly incorrectly cited). Mortality rates were adjusted according to change in HAQ score using an equation, but the source of the equation was not cited.

Estimates of costs

Resource use was based on HAQ score according to Taylor *et al.*[169]

Schering-Plough Ltd (infliximab)

Characteristics of starting population

The characteristics of starting population were based on GO-AFTER (a golimumab trial in patients who had inadequate response to TNF inhibitors): mean age 54, female 79%, baseline HAQ score of 1.61. The starting population was younger and had much lower baseline HAQ score than corresponding patients in the BSRBR. Baseline utility (EQ-5D and SF-6D) was imputed from baseline HAQ using simple linear regression (lower HAQ corresponding to higher utility). The consequence is that the estimated baseline utility may have been higher than it should be.

Treatment sequence

The model compared the five technologies against conventional DMARDs. It also compared each of the three TNF inhibitors followed by RTX against conventional DMARDs.

Estimates of clinical effectiveness – short term

Effectiveness of biologics was measured using ACR response, which was then mapped to EULAR response using an algorithm derived from GO-AFTER data.

Effectiveness data for biologics was obtained from a network meta-analysis of RCTs largely outside the scope. The validity of the network meta-analysis was questionable (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*). Effectiveness data for conventional DMARDs were obtained from EULAR response estimated by Brennan *et al.*[170] using regression analysis based on the BSRBR data. It appears that EULAR response for corresponding patients who switched to a second TNF inhibitor (rather than conventional DMARDs) was available from the same analysis, but these data were not used in the model. Instead, estimates of effectiveness for TNF inhibitors were taken from the MTC and thus the data for comparative effectiveness were obtained from different sources that may not be comparable.

Estimates of clinical effectiveness – long term

For patients receiving biologics, the base-case analysis assumed zero utility progression. A sensitivity analysis was carried out assuming that utility progression was equal to that observed in the BSRBR (by EULAR response), which suggests that utility worsens for EULAR good responders, is close to zero for moderate responders and improves marginally for non-responders.[171] This seems counterintuitive.

A further assumption was made that patients have the same radiological damage at the end of biologic treatment as at the start and therefore their ability to improve on further treatment was also retained. This was implemented in the model by holding age and disease duration constant for the time on biologic. The impact of this assumption is unclear and does not seem to have been explored in sensitivity analyses.

Mapping of effectiveness data to utility

For the base case, utility was estimated to be a function of EULAR response, treatment (on biologic treatment or not), health-state utility at time of treatment initiation, age, disease duration, number of previous DMARDs and gender according to an analysis of BSRBR data.[172]

Discontinuation rule and treatment duration

Withdrawal data for TNF inhibitors were taken from the BSRBR analysis of patients receiving a second TNF inhibitor.[173] All patients receiving biologics who did not achieve a moderate or good EULAR response were discontinued from treatment at 6 months. Treatment withdrawals were assumed to be the same for RTX and ABT. This assumption may overestimate the proportions of people who continue with these therapies although data are limited. For RTX, in the German registry [RABBIT (Rheumatoid Arthritis oBservation of BIologic Therapy); Stangfeld *et al.*[174]], 39% of people had no response after 6 months. However, at 12 months 68% of patients had gone on to receive a second infusion. What proportion of the remaining 32% goes on to receive a further infusion is not yet known. Further attrition with subsequent courses is likely but difficult to estimate.

Withdrawal data for conventional DMARDs was taken from Barton *et al.*[175]

Handling of adverse events and mortality

No impact of treatment on mortality was assumed in the model.

Estimates of costs

It was assumed that where possible the monitoring and administration for biologics and MTX were carried out concurrently. This seems appropriate. Two cost assumptions are presented for RTX based on a 6-month or 9-month dosing frequency. The 6-month dosing frequency was based on market research rather than on actual data from systematically collected data and may not be appropriate.

Vial optimisation was assumed in the base case. The assumptions are based on a questionnaire survey of rheumatology units (33% response rate). In many institutions vial sharing is achieved by central (pharmacy) preparation of infusions in advance of patient arrival. This can lead to drug wastage where patients are deemed not fit for infusion or fail to turn up. In any case, any savings from vial sharing are dwarfed by dose escalation.[176] In the cited systematic review, 44% of patients treated with IFX had the drug dose increased.

Roche (rituximab)

Characteristics of starting population

The starting population was based on the REFLEX trial:[124–126] mean age 52.4 years, 81% female, disease duration 11.9 years, prior DMARDs 2.5 (excluding MTX). Over half of the patients in

REFLEX[124–126] were recruited from the USA and thus the cohort does not reflect UK population/practice, as exemplified in the much younger age compared with the BSRBR cohort.

Treatment sequence

The treatment sequence did not contain AZA.

As mentioned before, while AZA has limited efficacy, this drug would still be tried in patients with resistant disease and thus should be used late in the sequence.

Estimates of clinical effectiveness – short term

For RTX and ABT, ACR response rates from TNF-IR ICs (based on trials of patients who had failed one or more TNF inhibitor) were used. For TNF inhibitors, ACR response rates from DMARD-IR MTC (based on trials of patients naive to TNF inhibitors) were firstly discounted by 30% and then used in the model. The validity for the DMARD-IR MTC was questionable (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*). The estimates of effectiveness for TNF inhibitors and RTX/ABT were therefore taken from a different set of analyses that are not comparable.

Estimates of clinical effectiveness – long term

Long-term HAQ progression for patients staying on treatment was set at zero (and also assumed to be zero for other biologics) according to observation from the LTE arm of the REFLEX trial.[139] A 6-monthly progression of 0.0225 was assumed for conventional DMARDs and 0.03 for palliative care. These were slightly lower than figures used in other manufacturer models.

Mapping of effectiveness data to utility

The HAQ scores were converted to EQ-5D scores according to the non-linear equation developed by Ducournau *et al.*[165] using data from TOC trials. An additional analysis that included age as a covariate in the non-linear model was also performed.

Discontinuation rule and treatment duration

Continuation of treatment (for all drugs) was subject to achieving an ACR20 or higher at the end of the first 6-month cycle. Subsequently, the same annual withdrawal rate (9.5%) for all biologics was assumed. This was based on Geborek *et al.*:[148] an average of two estimates for ETN (8%) and IFX (12%) used as the first biologic therapy. The same annual withdrawal rate (27%) was assumed for all traditional DMARDs. This was based on Bansback *et al.*,[146] which cited Wolfe[177] as the source. The data are likely to be outdated for some of the DMARDs.

Handling of adverse events and mortality

Adverse events were not included in the model.

Estimates of costs

Drug acquisition, administration and monitoring costs were estimated based on a 5-year average. This may not accurately reflect the costs of drugs with higher start-up costs.

Bristol-Myers Squibb Ltd (abatacept)

Characteristics of starting population

The characteristics of the starting population were based on the ATTAIN RCT.[130] Using data from a recent UK cohort (BSRBR[122]) might have been a more appropriate approach. Compared with the BSRBR data, patients in the ATTAIN trial[130] were on average slightly younger (58.0 years vs 53.4 years), and had a longer disease duration (9.0 years vs 12.2 years) and more patients were receiving glucocorticoids (44%–52% vs 70.2%). The mean HAQ score was slightly lower in the ATTAIN trial[130] than in BSRBR[122] data (1.8 vs 2.0) and the DAS28 score was slightly higher (6.5 vs 6.4).

Treatment sequence

It was assumed that a conventional DMARD is not likely to be used after a failure of the first TNF inhibitor. This is arguable and it is likely that at least a proportion of rheumatologists may seek to try drugs such as LEF, GST or CyA in this circumstance.

Penicillamine is included although it is used rarely today. The treatment sequences described, which were based on Barton *et al.*,[175] are credible.

Estimates of clinical effectiveness – short term

Clinical effectiveness in the first 6 months was estimated using HAQ scores. For RTX and ABT these were obtained from an MTC (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*). For TNF inhibitors the estimate was based on a BSRBR data analysis by the Decision Support Unit for NICE[178] and it used the adjusted result for switchers with long duration of second treatment (the report concluded that this is a good estimate for a year of treatment). For conventional DMARDs, data from early RA patients were used.[179–181] These data do not come from the population relevant to the scope (patients who failed a TNF inhibitor), but it was probably not possible to identify more relevant data.

Estimates of clinical effectiveness – long term

For long-term HAQ progression there were two sets of data: one versus RTX and one versus TNF inhibitors. For ABT there was a further HAQ reduction on treatment based on an analysis of ATTAIN and an extension of RTX trials[130,176] (–0.0729 and –0.013, respectively). For all other treatments (biologic drugs and conventional DMARDs) an annual increase in HAQ score of 0.012 was assumed based on an ERG STA report on RTX (calculation was actually based on non-biologic data).[183] It is unclear why only patients on ABT were assumed to further improve after the initial effect of the treatment, while all the other treatments are associated with deterioration.

Mapping of effectiveness data to utility

The algorithm mapping HAQ to utility was based on a conference abstract.[152] A linear equation (intercept 0.76, slope –0.28, female + 0.05) was used for that purpose.

Discontinuation rule and treatment duration

The treatment duration was based on data from ATTAIN LTE[119] for ABT (clinical study report 029). For all other treatments data, for first biologic use from Barton *et al.*[175] were utilised. As there were no data for ADA and RTX, an average for all biologics was assumed. These may not be directly applicable to the present decision problem.

The data used in the model differ from those in the BSRBR, but it is unclear if these parameters affect the results.

Discontinuation rates due to AEs in the first 6 months for ABT and RTX were based on a MTC (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*). For all other treatments, data from studies and reviews in TNF inhibitor naive patients were used.[179,184–188] The applicability of their results might be limited, although for conventional DMARDs probably no data in the relevant population were available. The proportion of patients discontinuing because of AEs was the lowest for ABT (2.3%) and ADA (2.8%) and was the highest for conventional DMARDs (12%–20%).

Handling of adverse events and mortality

The submission states that 'The event rates for ABT and RTX were derived from the mixed treatment comparison [please see comments]. The event rates for etanercept, adalimumab and

infliximab were derived from individual trials and the event rates for conventional DMARDs were based on the literature (as used in Chen *et al.*[179])'.

The utility loss due to AEs was based on data from an ERG STA report on erlotinib for relapsed non-small cell lung cancer.[189] Neutropenia and leucopenia were associated with a utility loss of 0.15 and all other AEs with a utility loss of 0.05. The applicability of these estimates to RA patients might be limited.

For mortality, a HAQ mortality HR of 1.33 (95% CI 1.10 to 1.61) was used based on Wolfe *et al.*[190]

Estimates of costs

The submission states that drug costs were based on the doses recommended in the drugs' summary of product characteristics. Drug treatment costs were taken from the Monthly Index of Medical Specialties (MIMS). The number of ABT vials used is assumed to be 2.85. This implies vial sharing. Currently less than 200 patients have been treated with ABT in the UK. Presently, it is unlikely that significant vial sharing can occur unless many more patients are treated. As dose wastage for IFX is assumed it would also be appropriate to model dose wastage with ABT.

Drug administration costs were based on Chen *et al.*[179] and an ERG STA on RTX.[183] Monitoring costs were based on Barton *et al.*[175] and Curtis.[191] These sources seem to be credible.

Hospitalisation resource use was based mainly on data from the NOAR Database (which included joint replacement).[197] Joint replacement surgery was included in the model separately and therefore it was deduced from the NOAR data assuming that two-thirds of RA hospitalisations are due to joint replacement (as stated in Pugner *et al.*[193]).

Time to joint replacement was assumed to be the same as in Barton *et al.*[175] and its impact on HAQ score was based on Wolfe and Zwillich.[194] The cost of joint replacement was assumed to be around £6,000.[195]

NHS Reference costs for 2007–8 were used for AEs [*as stated in the manufacturer's submission*[205]*; no citation provided*].

Discussion

A few common issues were identified in the critique of manufacturer models:

- starting population might not reflect UK population and practice
- validity and uncertainty in translating effectiveness measures into utility
- validity of ICs/MTCs carried out in trials of a heterogeneous population
- uncertainty in the relative effectiveness between individual TNF inhibitors and between these drugs and RTX/ABT
- uncertainty related to the effectiveness of conventional DMARDs
- uncertainty in long-term disease progression on various treatments
- different discontinuation rules
- different assumptions with regard to dosing interval or vial optimisation.

One particular challenge for this technology assessment/appraisal was an absence of RCTs for the three TNF inhibitors. It is the assessment group's view that evidence for technologies other than ABT and RTX is not appropriate for MTC or IC. Different approaches have been used by the assessment group and the manufacturers in this circumstance. The assessment group evaluated evidence from observational studies in detail in the absence of relevant RCTs for ADA, ETN and

TABLE 74 Data input and assumptions used in manufacturer models

	BRAM	Abbott	Wyeth Pharmaceuticals	Schering-Plough Ltd	Roche	Bristol-Myers Squibb Ltd
Baseline characteristics	Based on BSRBR:[163] mean age 58 years, 81% female, baseline mean HAQ 2.00	Based on BSRBR:[122] median age 58 years, 79% female, baseline mean HAQ 2.10	Based on TEMPO trial (a trial in TNF inhibitor naive patients): mean age 53 years, 74% female, baseline mean HAQ 1.60[a]	Based on GO-AFTER (golimumab trial in patients who had inadequate response to TNF inhibitors): mean age 54 years, 79% female, baseline mean HAQ 1.61	Based on REFLEX trial:[124–126] mean age 52.2 years, 81% female, baseline mean HAQ 1.88[b]	Based on the ATTAIN trial:[127–132] mean age 53.4 years, 77% female, baseline mean HAQ 1.80
Treatment sequence (after the failure of one TNF inhibitor)	Compared each of the five technologies against conventional DMARDs	Compared each of the five technologies against conventional DMARDs and a strategy of TNF inhibitor followed by RTX	Compared TNF inhibitor (as a class) with conventional DMARD and RTX	Compared each of the five technologies against conventional DMARDs. Also compared each of the three TNF inhibitors followed by RTX against conventional DMARDs	Compared each of the five technologies against conventional DMARDs	Compared various strategies of sequential use of two biologics
Estimates of clinical effectiveness – short term	Based on HAQ changes For biologics, data was obtained from the systematic review (where available – IFX was assumed to be the same as ETN due to lack of data); for DMARDs halved effectiveness in early RA was used due to lack of data for the relevant population	Based on ACR response rates mapped to HAQ changes Estimated ACR response rates were obtained from MTC of trials outside the scope (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*) Response rates were assumed to be equal for the three TNF inhibitors. Predicted response rates for the placebo/control arms were used as the response rates for conventional DMARDs ACR response rates were then mapped to HAQ changes using data from DE019 trial (ADA as first biologic therapy, outside scope)	Based on HAQ changes Short-term HAQ improvement for TNF inhibitors (as a class) was based on data from ReAct study,[96] an observational study of switching to ADA after failing a TNF inhibitor. Data from REFLEX trial[124–126] were used for RTX. ABT was not included in the model Assumed zero HAQ improvement for patients switched to conventional DMARDs according to BSRBR data[122]	Based on mapping between EULAR response and ACR response using an algorithm derived from the GO-AFTER trial data For biologics, effectiveness was estimated as odds ratios of ACR responses obtained from a network meta-analysis of RCTs largely outside the scope. The validity of the network meta-analysis was questionable (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*) The odds ratios were applied to baseline ACR response rates for conventional DMARDs, which were converted from EULAR response rates based on BSRBR data[173] The results were then converted back to EULAR responses using the aforementioned algorithm	Based on ACR response rates mapped to HAQ changes For RTX and ABT, ACR response rates from TNF-IR ICs (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*) were used. For TNF inhibitors, ACR response rates from DMARD-IR MTC were firstly discounted by 30% and then used in the model For conventional DMARDs, ACR response rates for placebo groups from TNF-IR ICs (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*) were used ACR responses were then mapped to HAQ changes using data from REFLEX trial[124–126]	Based on HAQ changes Short-term HAQ improvement was based on IC for ABT and RTX (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*), and on an analysis of BSRBR data by NICE Decision Support Unit[196] for TNF inhibitors For conventional DMARDs data from early RA trials were used[179–181]

DOI: 10.3310/hta15140

	BRAM	Abbott	Wyeth Pharmaceuticals	Schering-Plough Ltd	Roche	Bristol-Myers Squibb Ltd
Estimates of clinical effectiveness – long term	HAQ progression – base case: ■ Zero on biologics ■ 0.045/year on conventional DMARDs ■ 0.06/year on Pall	HAQ progression – base case: ■ 0.03/year on biologics (assumed the same as general population[168]) ■ 0.045/year on conventional DMARDs ■ 0.06/year for non-responders Zero HAQ progression on biologic treatment was explored in sensitivity analysis	HAQ progression – base case: ■ Zero on biologics ■ 0.15/year for medium term (6 months to 3 years) and 0.2/year for long term (beyond 3 years) on conventional DMARDs ■ 0.2/year on salvage therapy Sources include published literature (however, references seem to be incorrect)	Utility progression – base case: ■ Zero on biologics ■ Patients' age and disease duration were held constant for the time on biologic based on the assumption that no radiologic progression occurs during treatment ■ Utility progression on conventional DMARDs was estimated according to an annual HAQ progression of 0.042 based on Scott *et al.*[197]	HAQ progression – base case: ■ Zero on RTX (also assumed to be 0 for other biologics) according to observation from LTE of REFLEX trial[139] ■ 0.045/year on conventional DMARDs ■ 0.06/year for palliative care	HAQ progression – base case: ■ For ABT only there was a further HAQ reduction on treatment (0.073/year vs RTX and 0.013/year vs TNF inhibitors) based on an analysis of ATTAIN and an extension of RTX trials[130,182] ■ Increase in HAQ of 0.012/year for all other treatments (RTX, TNF inhibitors and conventional DMARDs) based on an ERG report on RTX[183]
Discontinuation rule and treatment duration	No formal withdrawal rule, but based on available data For long-term survival on treatment Weibull curves were fitted to the available data: ■ TNF inhibitors – from BSR submission ■ RTX – from REFLEX LTE[139] ■ ABT – from Bristol-Myers Squibb Ltd submission DMARDs – General Practice Research Database (GPRD) data	The minimal response required for continuation of treatment after the initial 6-month period is ACR50 Withdrawal rates used in the base-case analysis for TNF inhibitors are based on a shared frailty model previously developed by the Decision Support Unit using BSRBR data for patients receiving second TNF inhibitor. Withdrawal rates for ABT and RTX were assumed to be the same as for TNF inhibitors	Mentioned 'switching thresholds' based on relationship between HAQ and DAS28, but discontinuation rule was not clearly stated	Withdrawal data for TNF inhibitors (assumed the same for RTX and ABT) were taken from a BSRBR analysis.[173] Patients receiving biologics who do not achieve a moderate or good EULAR response were discontinued from treatment at 6 months Withdrawal data for conventional DMARDs were taken from Barton *et al.*[175] The utility rebound following treatment discontinuation was equal to the initial utility gain	Continuation of treatment (for all drugs) was subject to achieving an ACR20 or higher at the end of first 6 months. Subsequently, the same annual withdrawal rate (9.5%) for all biologics was assumed. This was based on Geborek *et al.*:[148] an average of two estimates for ETN (8%) and IFX (12%) used as the first biologic therapy Assumed the same annual withdrawal rate (27%) for all traditional DMARDs. This was based on Bansback *et al.*,[146] which cited Wolfe[177] as the source	The treatment duration was based on data from ATTAIN LTE[119] for ABT (clinical study report 029). For all other treatments data for first biologic use from Barton *et al.*[175] were utilised Discontinuation rates due to AEs in the first 6 months for ABT and RTX were based on an IC (see section *Critique of indirect comparisons and mixed-treatment comparisons included in manufacturers' submissions*). For all other treatments data from studies and reviews in TNF inhibitor-naive patients were used[179,184–188]

continued

TABLE 74 Data input and assumptions used in manufacturer models *(continued)*

	BRAM	Abbott	Wyeth Pharmaceuticals	Schering-Plough Ltd	Roche	Bristol-Myers Squibb Ltd
Mapping of effectiveness data to utility	Quadratic equation using dataset supplied by Hurst and reported in Hurst *et al.*[155] in the absence of any more recent dataset available to the assessment group	HAQ scores were converted to EQ-5D scores according to equations (linear and non-linear) developed by Ducournau *et al.*[165] using data from TOC trials. The non-linear equation was used for the base-case analysis, while the linear equation was examined in sensitivity analyses	HAQ scores were converted to EQ-5D scores according to a linear equation developed by Brennan *et al.*[150]	Utility was estimated to be a function of EULAR response, treatment (on biologic treatment or not), health-state utility at time of treatment initiation, age, disease duration, number of previous DMARDs and gender	HAQ scores were converted to EQ-5D scores according to the non-linear equation developed by Ducournau *et al.*[165] using data from TOC trials	HUI 3 utilities were calculated from the HAQ based on a conference abstract[153] EQ-5D utilities calculated from HAQ were used in a sensitivity analysis
AEs	Not incorporated into the model	Data on the occurrence of mild, moderate and serious AEs for ETN, IFX and LEF were estimated from Gebroek *et al.*[148] AEs for ADA, RTX and ABT were assumed to be the same as for ETN Rates of TB associated with each of the TNF inhibitors were based on data from the BSRBR[198]	Data were obtained from various literature sources (however, references seem to be wrong)	Not included in the model	Not included in the model	Occurrence of AEs was based on MTC (please see comments) Utility decrements for AEs based on ERG report on erlotinib for relapsed non-small cell lung cancer STA
Mortality	Basic mortality was taken from standard life tables. A RR (1.33) per unit HAQ was applied based on Wolfe *et al.*[199] For PSA a log-normal distribution was assumed (95% CI 1.10 to 1.61)	Assumed a reduction in mortality for patients receiving TNF inhibitors based on Jacobsson *et al.*[166] The reduction also applies to RTX or ABT	At baseline 1.63 times standard mortality from UK life tables. Adjusted based on HAQ Δ(mortality) = current mortality $\times$ [0.375 Δ(HAQ)]	No impact of treatment on mortality was assumed in the model	Mortality was adjusted according HAQ score based on Barton *et al.*,[175] which in turn was based on Wolfe *et al.*[199]	HAQ mortality HR based on Wolfe *et al.*[199]

	BRAM	Abbott	Wyeth Pharmaceuticals	Schering-Plough Ltd	Roche	Bristol-Myers Squibb Ltd
Drug costs and other costs	Costs are made up of drug and monitoring costs. A 'start-up' cost reflects higher dosage and additional monitoring, as appropriate for each treatment Unit costs were based on: ■ For tests and visits – values from Chen *et al.*[179] inflated to 2008 and from Curtis[191] For drugs – BNF 58 accessed online	Based on Monthly Index of Medical Specialties (July 2009) assuming an average patient weight of 70 kg Disease-related hospital costs (inpatient days and joint replacement procedures) were estimated based on HAQ band using data from the NOAR database[150]	Unit drug costs from BNF. Other costs from Curtis 2007 (Table 10 MS)	It was assumed that where possible the monitoring and administration for biologics and MTX was carried out concurrently. This seems appropriate. Two cost assumptions are presented for RTX based on a 6-month or 9-month dosing frequency. The 6-month dosing frequency was based on market research	Drug acquisition, administration and monitoring costs were estimated based on a 5-year average. This may not accurately reflect the costs of drugs with higher start-up costs	Drug costs were based on MIMS. Drug administration costs were based on Chen *et al.*[179] and an ERG report on RTX.[183] Monitoring costs were based on Barton *et al.*[175] and Curtis[191] Hospitalisation resource use was based mainly on data from the Barbieri study[192] The cost of joint replacement was assumed to be around £6000[195] NHS Reference costs for 2007–8 were used for AEs [*as stated in the manufacturer's submission;*[205] *no citation provided*]
Vial sharing	No vial sharing assumed	No vial sharing assumed	Not stated	Vial sharing for IFX was assumed	Not stated	The stated number of ABT vials used (2.85) implies vial sharing. See section *Bristol-Myers Squibb Ltd (abatacept)*
Time to retreatment for RTX	Base case: 8.7 months (based on Roche submission). Sensitivity analysis: 6 months	Base case: 9 months. Sensitivity analyses: 3, 6 or 12 months	Unclear	Both 6 months and 9 months were tested in base cases	Base case: 8.7 months (based on UK market research data, GfK HealthCare, January 2008; Roche data on file). Sensitivity analyses: 6, 7, 8 or 9 months	Base case: 6 months. Sensitivity analysis: 9 months

BNF, *British National Formulary*; BRAM, Birmingham Rheumatoid Arthritis Model; NOAR, Norfolk Arthritis Register; Pall, palliation.

a Percentage female and baseline HAQ score were not clearly stated. The figures shown were quoted in a related, published model[147] on which the industry submission was based.

b Figures were not clearly stated. The figures shown were obtained from Roche's previous submission for NICE Technology Appraisal 126, which was quoted in current submission.

DOI: 10.3310/hta15140

IFX, which is an unusual situation. The most appropriate data from either RCTs or observational studies for each of the technologies under assessment were then selected for economic modelling.

In order to conduct a valid IC, a network of RCTs that are comparable with respect to patient population and study design is needed. As stated above, no RCT conducted in a relevant patient population was found for the three TNF inhibitors. In order to perform ICs beyond ABT and RTX, one or more assumptions have to be made (as the manufacturers did):

- Assumption (1) – the effectiveness and safety of different TNF inhibitors are the same (e.g. evidence from trials of golimumab is applicable to the three TNF inhibitors under assessment).
- Assumption (2) – treatment effects are comparable between trials conducted in patients with different treatment history (DMARDs and biologics) and duration of RA, among other characteristics.

No evidence currently allows verification of assumption (1). To confirm or refute assumption (2) requires a systematic and comprehensive review far beyond the scope of this technology assessment/appraisal. Based on limited information provided in the MTCs included in the MSs, it appears substantial clinical, methodological and statistical heterogeneity exists among trials conducted in populations beyond the scope of this appraisal. The validity of analyses based on this assumption is thus questionable. It should therefore be borne in mind that potential uncertainties relating to these assumptions may not have been adequately reflected in the results of ICs/MTCs and the economic evaluations based on them.

Independent economic assessment

The assessment group's own independent analysis was carried out using the Birmingham Rheumatoid Arthritis Model (BRAM), which has been further updated to allow for a non-linear relationship between HAQ and utility. Additional coding has been added to the model to facilitate the use of PSA. This means putting a distribution around all parameters in the model. Unless there is a good reason to treat a parameter as fixed, some distribution has been used. Fixed parameters were: life tables, discount rates, treatment costs and times at which early withdrawal of treatment was assessed.

The BRAM is an individual sampling model. A large number of virtual patient histories are simulated with the accumulation of costs and QALYs. The basic model structure is shown in *Figure 93*. A complete description of the model follows here. A list of the assumptions in the model is given in *Appendix 15*.

Methods

Patients are assumed to follow a sequence of treatments. This involves starting a treatment, spending some time on that treatment, quitting a treatment if it is toxic or ineffective and starting the next treatment. The pattern is then repeated as long as active treatments are available. The final treatment in any strategy is palliation (Pall).

The HAQ DI (see *Appendix 1*) is used as the marker for disease severity. Scores on this scale range from 0 (best) to 3 (worst) in multiples of 0.125. Patients' HAQ scores are assumed to improve (decrease) on starting a treatment and this improvement is lost on quitting the treatment regardless of reason for quitting. While on treatment, a patient's condition is assumed to decline slowly over time. This is modelled by occasional increases of 0.125 in HAQ score. The mean time between such increases in HAQ is allowed to vary by treatment; see *Figure 94* for a possible HAQ

trajectory. In the reference case analysis, HAQ is assumed to remain constant while a patient is successfully treated with a biological agent: this is modelled by a very large mean time to increase in HAQ.

Strategies to be compared

The current appraisal is concerned solely with the decision to be made at the point of failure of a first TNF inhibitor. Accordingly, the starting population consists of patients who have reached that point in a sequence of treatments. *Table 75* shows the treatment sequences compared in this appraisal.

Note, that previous versions of the BRAM used a starting population of DMARD-naive patients, and generated a range of different decision populations within the model. Strategies compared also allowed different choices of treatment options depending on the toxicity of previous treatments. While the coding to allow this flexibility remains within the model, such flexibility is not required within the present appraisal.

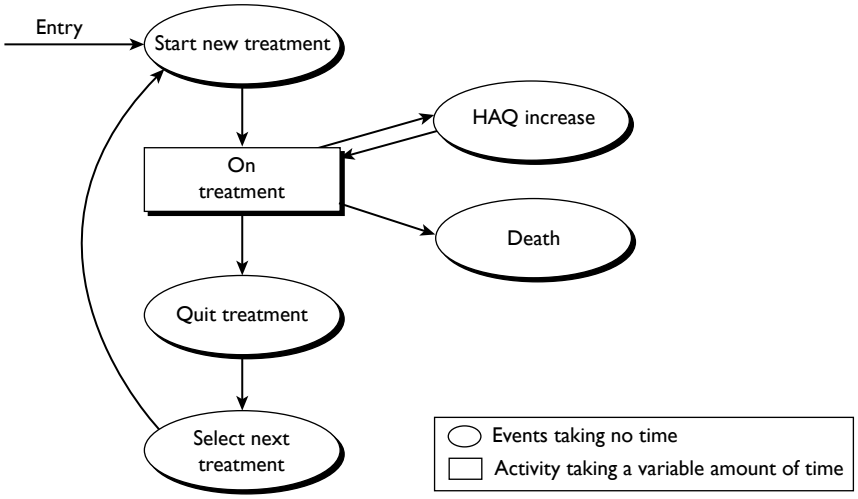


FIGURE 93 Basic structure of the BRAM individual sampling model.

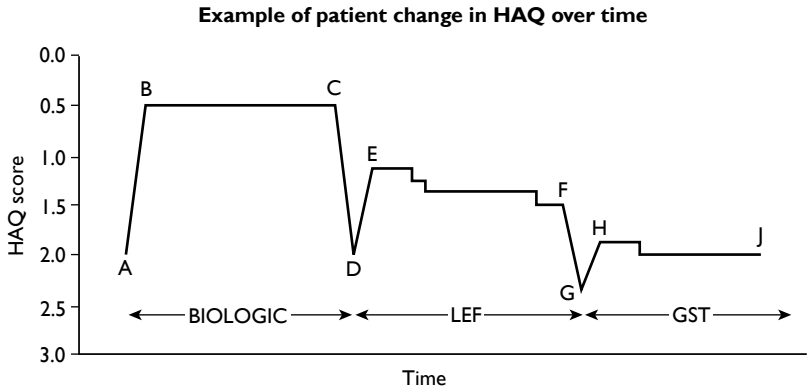


FIGURE 94 Possible trajectory of HAQ score over time. Initial improvement on a biological agent (AB) is lost on quitting the treatment (CD). A smaller improvement (DE) on starting LEF is similarly lost on quitting (FG) and followed by a gain (GH) on starting GST. In this case the patient dies of other causes (J) while still responding to GST. There is a gradual deterioration in HAQ from E to F and from H to J, but not from B to C in the reference case analysis. In some cases, the time spent on a conventional DMARD is not long enough for any deterioration in HAQ to occur.

TABLE 75 Treatment sequences compared in the BRAM for this appraisal

Strategy name	ADA	ETN	IFX	RTX	ABT	DMARDs
First	ADA	ETN	IFX	RTX	ABT	LEF
Second	LEF	LEF	LEF	LEF	LEF	GST
Third	GST	GST	GST	GST	GST	CyA
Fourth	CyA	CyA	CyA	CyA	CyA	AZA
Fifth	AZA	AZA	AZA	AZA	AZA	Pall
Sixth	Pall	Pall	Pall	Pall	Pall	

All biologics are assumed to be taken in combination with MTX.

The choice of DMARDs following biologic therapy has been made in line with expected practice and excludes any DMARDs that are likely to have been used before biologic therapy.

Data used in the Birmingham Rheumatoid Arthritis Model

What follows is a detailed description of the data and sources thereof. Updated literature reviews have been used wherever possible.

Initial patient data

Tables 76 and *77* show the information about the initial population. As stated earlier, the initial population is a population immediately following failure of a first TNF inhibitor. The values are based on the BSRBR submission to NICE.[163]

Starting treatments

As in the previous version of the BRAM, the change in HAQ on starting a new DMARD is sampled on an individual basis and takes the form of a multiplier applied to the HAQ score on starting treatment. This multiplier is sampled from a beta distribution. The method used to estimate the parameters of the beta distribution is the same as in a previous report.[179]

To illustrate the method, consider the calculations used in the previous report for LEF. The data available were baseline HAQ [mean 1.03, standard deviation (SD) 0.62] and HAQ improvement [mean 0.48, (SD) 0.5].[1,100] An EXCEL spreadsheet was set up to create a starting population of 10,000 virtual patients with HAQ scores drawn from a normal distribution with mean and SD supplied by the user. Each generated HAQ score was converted to the nearest legitimate value (multiples of 0.125 in the range 0–3). The parameters supplied were adjusted to compensate for the effect of this conversion, so that the mean and SD of the population generated corresponded to the data. In this case, this involved adjusting the mean of the underlying distribution to 1.01 and the SD to 0.66. The sample mean and SD then agreed with the data.

A beta distribution was found to match the given mean and SD for HAQ improvement. In this case the parameters were $a = 0.57$ and $b = 0.65$. *Figure 95* shows the simulated population in this case. Each square within the graph represents a possible pair of values of starting HAQ and HAQ on treatment: the darker the square, the larger the number of simulated patients with that pair of HAQ values. It can be seen that there was a high proportion of patients with equal HAQ on treatment compared with before treatment. In this example, the sampled population contained a large number of zero initial HAQ values. These are omitted from the graphs, but included in the calculations relating to HAQ improvement.

In the current report, for biologic DMARDs, the parameters have been re-estimated using the best available data for use immediately after a first TNF inhibitor. For conventional DMARDs to be used after biologics, the only available data were from trials in early RA. The effectiveness was halved for use in late RA.

TABLE 76 Initial age and gender distribution

	Age (years)							
Gender	**15–24**	**25–34**	**35–44**	**45–54**	**55–64**	**65–74**	**75–84**	**Total**
Male	0.0	0.4	1.9	5.2	6.5	3.8	1.2	19
Female	0.1	1.5	8.2	22.1	27.7	16.3	5.1	81

TABLE 77 Starting distribution of HAQ scores

HAQ	**0.125**	**0.25**	**0.375**	**0.5**	**0.625**	**0.75**	**0.875**	**1**
%	0.0	0.1	0.2	0.5	0.7	1.2	1.7	2.2
HAQ	**1.125**	**1.25**	**1.375**	**1.5**	**1.625**	**1.75**	**1.875**	**2**
%	2.9	3.6	4.3	5.1	5.8	6.6	7.2	7.7
HAQ	**2.125**	**2.25**	**2.375**	**2.5**	**2.625**	**2.75**	**2.875**	**3**
%	8.1	8.4	8.3	8.0	7.1	5.9	3.7	0.7

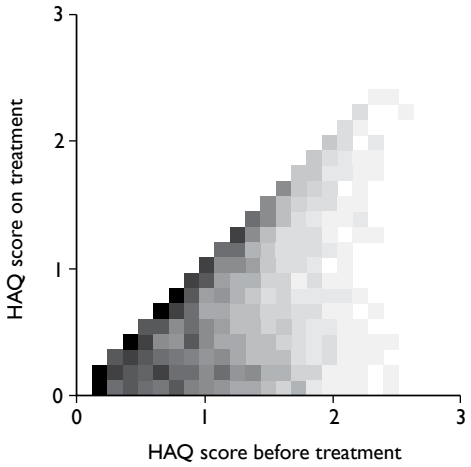


FIGURE 95 Modelled distribution of HAQ score change on starting leflunomide (from previous report).[177]

When a patient starts a new treatment in the model, a random number is drawn to determine the HAQ improvement for that patient. Consider, for example, a patient about to start LEF with a HAQ score of 2 and suppose that the random number drawn is 0.5. The value of 0.5 indicates that the improvement multiplier should be at the median of the relevant distribution. In the case of LEF, using the values from *Table 78*, the median is 0.358 so the HAQ should improve by $0.358 \times 2 = 0.716$. However, because HAQ is measured on a discrete scale, the improvement must be rounded to the nearest multiple of 0.125, which in this case is 0.75. The HAQ score on treatment would then be $2 - 0.75 = 1.25$, and the 0.75 improvement (reduction) would be lost on quitting treatment. Had the starting HAQ score been 1, the improvement would have been 0.375 to give a HAQ on treatment of 0.625.

Table 78 shows the point estimates for the parameters of the beta distributions used. However, these values are not known with certainty, so some variation must be included in the PSA. In the absence of any obvious way of measuring the uncertainty around the parameters, an assumption was made that each could be independently sampled from a normal distribution with an SD equal to 0.1 times the point estimate. This is still likely to underestimate the uncertainty in these

TABLE 78 Beta distributions for HAQ multipliers (point estimates)

Treatment	*a*	*b*	Mean	HAQ improvement on starting treatment/baseline HAQ; source
ADA	0.32	0.92	0.26	0.48/1.85; Bombardieri 2007[95,96]
ETN	0.21	0.75	0.22	0.35/1.60; Bingham 2009[104]
IFX	0.21	0.75	0.22	Assume same as ETN
RTX	0.20	0.75	0.21	0.40/1.90; REFLEX[124–126]
ABT	0.33	0.85	0.28	0.50/1.80; ATTAIN[127–132]
LEF	0.285	0.935	0.23	Effectiveness halved from values used in previous report[179]
GST	0.225	0.925	0.20	
CyA	0.065	0.325	0.17	
AZA	0.10	0.90	0.10	

For PSA, the values *a* and *b* are drawn from normal distributions with SD 0.1 times the point estimate (see text).

parameters, but is preferable to using fixed values. Note that, although the same point estimates have been used for ETN and IFX, separate and independent samples have been used for the two drugs in the PSA. This principle has been applied throughout the model. In such cases, it is not known in which direction the difference between the treatments should be, but it is not a reasonable assumption that the treatments should take identical values.

Added in response to consultees' comments: the values here give LEF a higher immediate effectiveness than any of the biologics. This is offset in part by the assumption described below about changes in HAQ score while on treatment. However, it is stressed that these values are not being used for a comparison in which the biologic treatments replace LEF in a sequence of treatments. Additional scenario analyses have been added to consider alternative assumptions.

Time on treatments The model allows for two stages of early quitting of treatment. For conventional DMARDs, this facility has been used with parameters preserved from Chen *et al.*[179] For TNF inhibitors and ABT, a single stage of early quitting has been included in line with available data, while for RTX no early quitting can be allowed, because it is necessary to model the full costs of each cycle of treatment. The values used are in *Table 79*. For long-term survival on treatment, Weibull curves were fitted to the available data.

In the form used, a random variable X has a Weibull distribution with shape parameter a and scale parameter b if:

$$\left(\frac{X}{b}\right)^a$$

has an exponential distribution with unit mean. If $a = 1$ the Weibull reduces to the exponential distribution with mean b; in any case b is the time until:

$$\frac{1}{e} \approx 37\%$$

of the original population remains. If $a < 1$ then the hazard decreases with time; if $a > 1$ the hazard increases. The values used are shown in *Table 80*. For convenience, the mean of the distribution is also shown for the point estimates of the parameters.

TABLE 79 Probability of early quitting of treatment

Treatment	Parameter	Point estimate (%)	Distribution	Source
ADA	Withdrawal at 12 weeks	9.9	Beta (89, 810)	Bombardieri 2007[95,96]
	Toxicity if above	56.2	Beta (50, 39)	
ETN	Withdrawal at 13 weeks	5.2	Beta (21, 385)	Bingham 2009[104] and Buch 2005[99]
	Toxicity if above	16.7	Beta (2, 10)	Bingham 2009[104]
IFX	Withdrawal at 16 weeks	23	Beta (3, 10)	OPPOSITE[133]
	Toxicity if above	66.7	Beta (2, 1)	
RTX	No early withdrawal (see text)			
ABT	Withdrawal at 6 months	13.6	Beta (35, 223)	ATTAIN[127–132]
	Toxicity if above	25.7	Beta (9, 26)	
LEF	Withdrawal at 6 weeks	13	Beta (13, 87)	Geborek 2002[148]
	Withdrawal 6–24 weeks	30	Beta (30, 70)	
	Toxicity if above	33.3	Beta (10, 20)	
GST	Withdrawal at 6 weeks	14	Beta (10, 62)	Hamilton 2001[1101]
	Withdrawal 6–24 weeks	27.1	Beta (19.5, 52.5)	
	Toxicity if above	66.7	Beta (6.5, 13)	
CyA	Withdrawal at 6 weeks	8	Beta (16, 184)	Yocum 2000[202]
	Withdrawal 6–24 weeks	24	Beta (48, 152)	
	Toxicity if above	50	Beta (24, 24)	Marra 2001[203]
AZA	Withdrawal at 6 weeks	15	Beta (15, 85)	Willkens 1995[204]
	Withdrawal 6–24 weeks	25	Beta (25, 75)	
	Toxicity if above	50	Beta (12.5, 12.5)	

TABLE 80 Times to quitting treatments

Treatment	*a*	95% CI	*b* (years)	95% CI	Mean (years)	Source
TNF inhibitors	0.701	0.634 to 0.768	3.211	3.022 to 3.412	4.06	BSRBR submission[123]
RTX	0.474	0.403 to 0.545	5.1	3.742 to 6.951	11.31	REFLEX LTE[139]
ABT	0.81	0.734 to 0.886	5.49	5.166 to 5.834	6.17	BMS submission[205]
LEF	1	0.905 to 1.095	5.98	5.627 to 6.355	5.98	GPRD database[206]
GST	0.48	0.434 to 0.526	1.81	1.703 to 1.923	3.91	
CyA	0.5	0.452 to 0.548	4.35	4.094 to 4.623	8.70	
AZA	0.39	0.353 to 0.427	4.35	4.094 to 4.623	15.53	

GPRD, General Practice Research Database; Normal distributions used for parameter: *a*, log-normal for parameter; *b*, Standard errors for TNF inhibitors and RTX estimated from data. For other treatments, the same proportional variability as for TNF inhibitors has been assumed. Mean time on treatment based on the point estimate of the parameters.

For TNF inhibitors, the same principle as for initial effectiveness has been applied: independent samples were drawn each time from the same distribution. For RTX, the time sampled is then taken up to the nearest multiple of the assumed time between treatment cycles.

Details of the implementation are as follows. For conventional DMARDs, the survival time is assumed to follow a distribution of the type shown in *Figure 96*, which is based on the data for LEF. The first step represents cessation of treatment after 6 weeks, which is assumed to be for toxicity. The second step represents cessation between 6 and 24 weeks after starting treatment, which could be for toxicity or inefficacy. At each appropriate stage in the running of the model,

two variables, $u1$ and $u2$, are each drawn from a uniform distribution between 0 and 1. *Figure 97* shows how these numbers are used. The value of $u1$ is first used in the beta distribution to determine the HAQ improvement described earlier. Then $u2$ is used to determine the time on treatment.

In implementation, critical values are calculated each time the population parameters are sampled for each treatment, so that the areas of the four zones in *Figure 97* correspond to the probabilities sampled from the distributions indicated in *Table 79*. Then, for each individual, the values of $u1$ and $u2$ are compared with those critical values in the following ways:

- If $u2$ is below its lower critical value, then the individual is in Zone A, and withdraws because of toxicity after 6 weeks.
- Otherwise, $u1$ is compared with its critical value. If $u1$ is below the critical value, then the individual is in Zone B, and withdraws because of ineffectiveness after 24 weeks.
- Otherwise, $u2$ is compared with its higher critical value. If $u2$ is below this value, then the individual is in Zone C, and withdraws because of toxicity after 24 weeks.
- Otherwise, the individual is in Zone D, and remains on treatment beyond 24 weeks. The value of $u2$ is converted to a value from the appropriate Weibull distribution to determine the time on treatment.

For TNF inhibitors and ABT, the 6-week quitting was not used and the time shown in *Table 79* was used in place of the 24-week limit used for conventional DMARDs. The implication of this is that for all modelled treatments except RTX, those individuals with the lowest HAQ improvement on starting treatment all quit early.

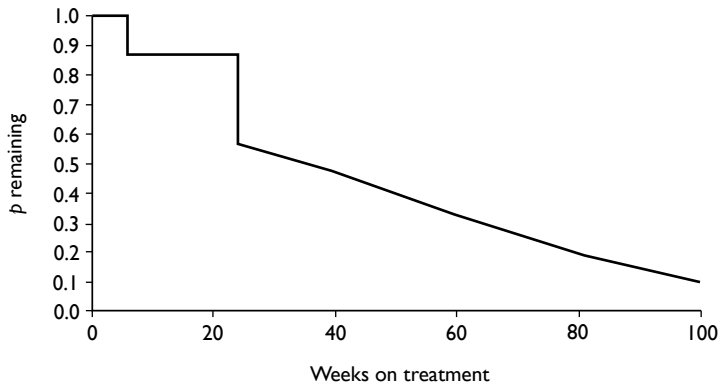


FIGURE 96 Survival time on a treatment (based on LEF data).

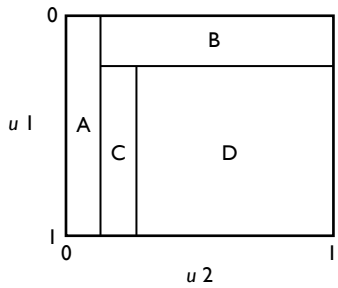


FIGURE 97 Early cessation of treatment. The four zones represent the following: A, withdrawal within 6 weeks (assumed due to toxicity); B, withdrawal between 6 and 24 weeks for inefficacy; C, withdrawal between 6 and 24 weeks for toxicity; and D, remaining on the treatment after 24 weeks.

Health Assessment Questionnaire changes on treatment In the reference case analysis, it is assumed that HAQ remains constant while on any biologic treatment. Mean rates of HAQ increase of 0.045/year on conventional DMARDs and 0.06/year on Pall are modelled as mean times to increase (by 0.125) of 2.7 years and 2 years, respectively. In the PSA these times are sampled from normal distributions with SDs 0.27 years and 0.2 years, respectively. Again, the times for the conventional DMARDs are sampled independently each time.

Costs Costs are made up of drug costs plus monitoring costs. As in previous versions, the model includes an annual usage cost for each treatment, together with a 'start-up' cost reflecting higher dosage and additional monitoring early in treatment, as appropriate for each treatment. *Table 81* shows the unit costs for tests and visits and *Table 82* the unit costs for drugs, leading to annual costs in *Table 83*.

An administration cost of £141.83 is assumed for each dose of IFX, RTX, and ABT. This figure is inflated from the figure of £124.00 used in earlier versions of the BRAM. Annual administration costs are shown in *Table 84*. Monitoring assumptions for conventional DMARDs are shown in *Table 85*. Annual cost for tests and administration are shown in *Tables 86* and *87*, respectively. It is assumed that monitoring for biologic therapies is included within the monitoring for MTX or administration costs, so no additional monitoring cost is included for these. Combining

TABLE 81 Unit costs for tests and visits

Test	Cost (£)	Source
FBC	4.55	Values from Chen 2006[179] inflated to 2008 prices using the Hospital and Community Health Services inflation index (Curtis 2008)[207]
ESR	3.51	
BCP	4.39	
CXR	17.82	
Urinalysis	0.09	
Visit		
GP	36	Curtis 2008[207]
Hospital outpatient	71	
Specialist nurse visit	35.50	Assumed half of outpatient visit
Administration of infusion	141.83	Chen 2006[179] inflated to 2008 prices

BCP, biochemical profile; CXR, chest X-ray; FBC, full blood count.

TABLE 82 Unit costs for drugs

Treatment	Cost	Assumptions
ADA	£357.50 per dose	26 doses per year
ETN	£178.75 per dose	52 doses of 50 mg per year
INF	£419.62 per vial	70-kg patient; drug wastage
RTX	£873.15 per 500-mg vial	Dosage of two × 1,000 mg every 8.7 months in base case
ABT	£242.17 per 250 mg	750 mg every 4 weeks
MTX	11.7p per tablet	15 mg per week
LEF	£1.70 per day	20 mg per day
GST	£11.23 per dose	50-mg ampoule administered at GP visit
CyA	£5.37 per day	225 mg per day
AZA	40.3p per day	150 mg per day

Source: *British National Formulary* 58 accessed online.

TABLE 83 Drug costs: first and subsequent years

Treatment	Cost (£) (steady-state yearly)	Cost (£) (additional in first year)	Assumptions
ADA	9,295	0	Twenty-six doses per year
ETN	9,295	0	Fifty-two doses of 50 mg per year
INF	7,553.16	1,258.86	Six doses per year; one additional dose in first year; three vials per dose
RTX	4,817.38	0	Each course is four 500-mg vials; multiply by 12/8.7 for annual cost
ABT	9,444.63	726.51	Thirteen doses of 750 mg = 39 times unit cost; one additional dose in first year
MTX	36.50	0	Six tablets per week for 52 weeks
LEF	620.50	0	365 times daily cost
GST	134.76	224.60	Steady-state 12 doses per year; additional 20 doses in first year
CyA	1,960.05	0	365 times daily cost
AZA	147.10	0	365 times daily cost

TABLE 84 Administration costs: first and subsequent years

Treatment	Cost (£) (steady-state yearly)	Cost (£) (additional in first year)	Assumptions
ADA	0	106.50	Three visits to nurse specialist
ETN	0	106.50	Three visits to nurse specialist
INF	850.98	141.83	Six doses per year; one additional dose in first year
RTX	391.26	0	Two infusions per course; multiply by 12/8.7 for annual cost
ABT	1,843.79	141.83	Thirteen infusions per year; one additional infusion in first year
MTX	0	0	
LEF	0	0	
GST	432	720	Steady-state 12 doses per year; additional 20 doses in first year; GP visit for each dose
CyA	0	0	
AZA	0	0	

the monitoring assumptions with the unit costs then leads to start-up and annual usage costs as shown in *Table 88*. Note, that as these costings are based on fixed prices and monitoring rules, rather than measured resource use, the prices are not varied in the PSA. All costs were discounted at 3.5% per annum from the start of the model.

Costs for hospitalisation and joint replacement are estimated by a cost per unit HAQ score. In the base-case analysis, this was set at £1,120.00 per unit HAQ. This was inflated from the previous figure of £860.00 per unit included in previous versions of the BRAM. Scenario analysis includes various alternative costings here based on industry submissions.

Mortality Basic mortality was taken from standard life tables. A RR per unit HAQ was applied. The point estimate for this RR was set to 1.33, sampling in the PSA from a log-normal distribution with 95% CI (1.10 to 1.61).

TABLE 85 Monitoring assumptions

Treatment	Pre-treatment	On treatment
MTX	FBC, ESR, BCP, CXR	FBC and BCP every 2 weeks for 4 months then monthly
LEF	FBC, ESR, BCP, urinalysis	FBC every 2 weeks for 6 months, every 8 weeks thereafter. BCP monthly for 6 months, every 8 weeks thereafter
GST	FBC, ESR, BCP, urinalysis	FBC and BCP, urinalysis every week for up to 21 injections, then every 2 weeks for 3 months, then every 3 weeks for 3 months, then monthly. Treatment given by i.m. injections
CyA	FBC, 2×BCP, ESR, urinalysis	FBC and BCP every 2 weeks for 4 months, then BCP monthly
AZA	FBC, ESR, BCP	FBC and BCP weekly for 6 weeks, then every 2 weeks for three visits, then monthly
Pall		Outpatient visit every 3 months

BCP, biochemical profile; CXR, chest X-ray; FBC, full blood count; i.m., intramuscular.

TABLE 86 Test costs: first and subsequent years

Treatment	Cost (£) (pre-treatment)	Cost (£) (steady-state yearly)	Cost (£) (additional in first year)
MTX	30.27	107.28	35.76
LEF	12.54	53.64	54.12
GST	12.54	108.36	180.60
CyA	16.93	52.68	53.96
AZA	12.45	107.28	53.64

TABLE 87 Test administration costs: first and subsequent years

Treatment	Cost (£) (pre-treatment)	Cost (£) (steady-state yearly)	Cost (£) (additional in first year)
MTX	71	852	284
LEF	71	426	639
GST	71	852	1,420
CyA	71	852	142
AZA	71	852	426

TABLE 88 Treatment costs

Treatment	Start-up (£)	Annual use (£)
ADA	527.53	10,290.78
ETN	527.53	10,290.78
IFX	1,821.72	9,399.92
RTX	421.03	6,204.42
ABT	1,289.37	12,284.20
LEF	776.66	1,100.14
GST	2,628.74	1,527.12
CyA	283.89	2,864.73
AZA	563.09	1,106.38
Pall	0.00	284.00

Quality of life scores In the reference case analysis, a quadratic equation was used to relate HAQ score to QoL score. This was of the form $QoL = a - b_1 \times HAQ - b_2 \times HAQ^2$ where the coefficients are shown in *Table 89*. It is noted that this equation gives negative values (indicating a state worse than death) for high HAQ scores. While this reflects the fact that individual patients in the dataset used to generate the equation gave EQ-5D responses that map to scores below zero on the standard UK tariff, it is acknowledged that the use of negative QoL scores is controversial. Accordingly, coding was added to allow such scores to be adjusted to zero in the model. This coding was used in scenario analysis.

It was assumed that start and end effects could be modelled as one-off deductions proportional to the change in QoL score. The multiplier was set to a base-case value of 0.2 (years), sampled from a normal distribution with an SD 0.02 (separately for start and end).

Accumulated QALYs were discounted at 3.5% per annum from the starting point of the model.

Results

When an individual sampling model is run with a fixed parameter set, it must be run with a large number of patients to produce a precise estimate of the population mean cost and QALY differences between strategies. When such a model is run using PSA, the aim is to produce a distribution for the population outcomes that reflects the parameter uncertainty. This is done by sampling repeatedly from the joint distribution of parameters, and then for any parameter set, sampling a sufficient number of individuals.

Figure 98 shows the overall design of such a model run.

Note that a new set of patients is sampled for each parameter set, but the same patients are run through each of the possible strategies. Trial runs were made with different numbers of patients per parameter set. At fewer than 2,000 patients, the distribution of points in the cost-effectiveness

TABLE 89 Coefficients in HAQ to QoL equation

Coefficient	Point estimate	95% CI
a	0.804	0.711 to 0.897
b_1	0.203	0.054 to 0.351
b_2	0.045	–0.007 to 0.096

Source: Birmingham analysis of dataset from Hurst. Note that the coefficient b_2 takes a negative value in approximately 9% of model replications. However, the positive value of b_1 ensures that QoL decreases with increasing HAQ.

Parameter set 1: QoL = 0.7688 – 0.1723 × HAQ – 0.0506 × HAQ^2, etc.	
	Patient 1.1: Female, starting age 45.0947, starting HAQ 2.875
	Patient 1.2: Female, starting age 51.2780, starting HAQ 2.75
	Repeat up to patient 1.*M*
Parameter set 2: QoL = 0.8209 – 0.2087 × HAQ – 0.0359 × HAQ^2, etc.	
	Patient 2.1: Female, starting age 50.6852, starting HAQ 2.625
	Patient 2.2: Female, starting age 59.4641, starting HAQ 1.625
	Repeat up to patient 2.*M*
Repeat up to parameter set *N*.	

FIGURE 98 Running an individual sampling model under PSA.

plane became visibly wider. For safety, we used 5,000 patients per parameter set. For the reference case analysis, 2,000 parameter sets were sampled from the parameter distributions as described in the previous section. For each parameter set, 5,000 individual patient attributes were sampled and these patients were run through each of the six strategies defined in *Table 75*.

Reference case

The discounted lifetime costs and QALYs for each patient were calculated and the mean results for each parameter set output. The overall mean of these results forms the reference case estimate for the mean cost and QALY of each strategy: the 2.5 and 97.5 percentiles give the limits of the 95% credible interval. Note that these percentiles are likely to come from different parameter sets not just between strategies, but also for costs and QALYs for any particular strategy. These results are shown in *Table 90*. In each case, the lower credible limit for QALYs is negative, reflecting the use of an equation that allowed negative QoL scores; see the *Scenario analysis* for the effect of changing this assumption.

Incremental results were obtained by subtraction for each parameter set, thus producing a sample of 2,000 points from the incremental cost-effectiveness distribution between any pair of strategies. Again, the 95% credible interval can be found for cost and QALY differences: note that, although the mean results can be inferred from *Table 90* (subject to rounding effects), the relevant percentiles cannot. The results are shown in *Table 91*, which shows all the pair-wise comparisons. Scatter plots for the comparisons between the biologic strategies and conventional DMARDs alone are shown in *Figure 99*, together with the CEACs for these five comparisons: the remaining scatter plots are shown in *Appendix 13*.

Similar remarks apply to the ICER, which is found by dividing the difference in mean cost by the difference in mean QALY. Finally, the proportion of model replications for each biologic strategy appears cost-effective compared with any other is shown, using a threshold ICER of £20,000/QALY and £30,000/QALY. These results are shown in *Table 92*.

Scenario analysis

A number of different scenarios have been run. Details of each scenario and the results are given in *Appendix 14*, and a summary is provided in *Tables 93–95*. It should be noted that, although it is always possible to give a result based on the mean of the probabilistic analysis, the results for comparison between TNF inhibitors almost invariably are from a distribution covering all four quadrants of the cost-effectiveness plane, and thus the mean results are subject to enormous uncertainty in that case. The sole exception to this is the scenario 'Vary time on TNF inhibitors'.

Summary of model results

The reference case model results show similar costs and QALYs for the TNF inhibitors, with somewhat lower costs and QALYs for RTX and higher costs and QALYs for ABT. Compared

TABLE 90 Results for single strategies in reference case analysis

Treatment	Mean cost (£)	95% credible interval		Mean QALY	95% credible interval	
ADA	74,800	68,800	81,000	2.89	–2.12	7.87
ETN	75,100	68,700	81,500	2.80	–2.21	7.84
IFX	73,000	66,100	79,700	2.80	–2.24	7.82
RTX	69,400	62,700	76,400	3.10	–1.78	7.95
ABT	93,000	86,200	100,100	3.28	–1.46	8.05
DMARDs	49,000	43,300	54,900	2.13	–3.27	7.46

TABLE 91 Differences between strategies in reference case analysis

Comparison	Diff cost (£)	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,800	24,100	27,500	0.75	0.33	1.23
ETN–DMARDs	26,100	24,200	27,900	0.67	0.30	1.10
IFX–DMARDs	24,000	19,500	26,800	0.67	0.29	1.12
RTX–DMARDs	20,400	17,500	23,200	0.96	0.41	1.61
ABT–DMARDs	44,000	41,300	46,700	1.15	0.52	1.88
ADA–RTX	5,400	2,200	8,700	–0.21	–0.52	0.03
ETN–RTX	5,700	2,400	9,100	–0.29	–0.63	–0.04
IFX–RTX	3,600	–1,600	7,600	–0.30	–0.62	–0.05
ABT–RTX	23,600	19,800	27,400	0.18	–0.10	0.50
ADA–ABT	–18,200	–21,300	–15,200	–0.39	–0.77	–0.12
ETN–ABT	–18,000	–21,200	–14,600	–0.47	–0.88	–0.17
IFX–ABT	–20,000	–25,100	–16,200	–0.48	–0.88	–0.17
ADA–ETN	–300	–2,800	2,100	0.08	–0.09	0.29
ADA–IFX	1,800	–1,400	6,500	0.09	–0.10	0.29
ETN–IFX	2,000	–1,200	6,800	0.00	–0.17	0.19

Diff, difference.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

TABLE 92 Incremental cost-effectiveness ratios for reference case analysis

				Proportion of cases cost-effective at	
Comparison	ICER (£/QALY)	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	34,300	20,900	79,100	0.02	0.30
ETN–DMARDs	38,900	23,500	89,000	0.00	0.17
IFX–DMARDs	36,100	21,200	82,000	0.02	0.24
RTX–DMARDs	21,100	12,800	49,700	0.40	0.84
ABT–DMARDs	38,400	23,000	84,700	0.00	0.17
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	130,600	47,900	RTX	0.00	0.00
ADA–ABT	*46,400*	*23,100*	*152,100*	0.99	0.90
ETN–ABT	*37,800*	*20,100*	*102,300*	0.98	0.77
IFX–ABT	*41,700*	*22,000*	*113,500*	0.99	0.84
ADA–ETN	ADA	Not meaningful		0.84	0.84
ADA–IFX	20,500	Not meaningful		0.50	0.61
ETN–IFX	456,700	Not meaningful		0.20	0.24

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one-half of the plane.

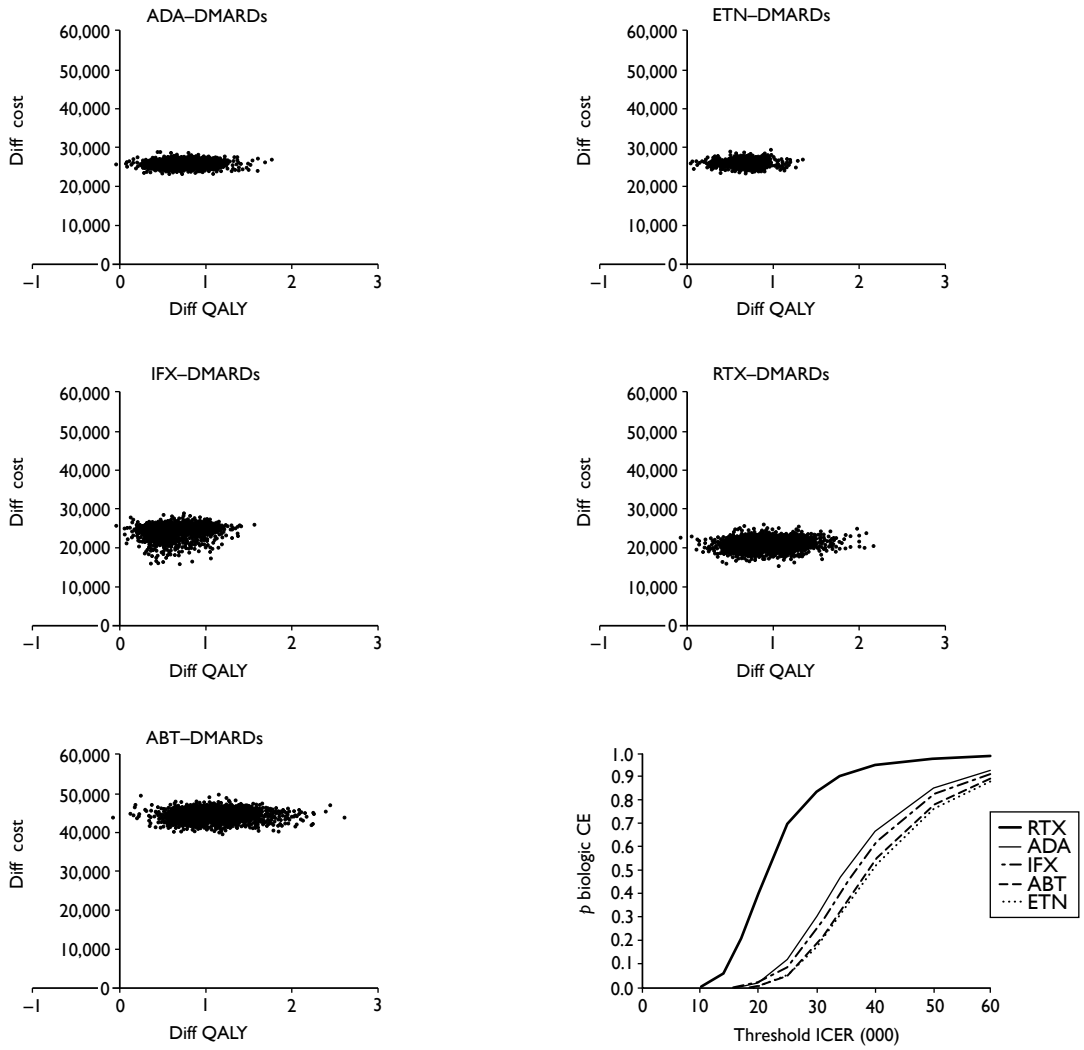


FIGURE 99 Cost-effectiveness scatter plots for main comparisons in the reference case. Diff, difference.

TABLE 93 Results from scenario analysis: comparisons against DMARDs strategy (ICER in £/QALY)

Scenario	ADA–DMARDs	ETN–DMARDs	IFX–DMARDs	RTX–DMARDs	ABT–DMARDs
Reference	34,300	38,900	36,100	21,100	38,400
Vary time on TNF inhibitors	34,300	38,400	37,700	21,200	38,500
Same time on all biologics	34,400	38,700	35,900	21,100	39,500
RTX cycle time 6 months	34,300	38,900	35,900	32,600	38,400
RTX cycle time 11.6 months	34,200	38,800	35,900	11,400	38,400
Poor late DMARDs	28,100	31,100	28,800	16,300	32,100
HAQ change on biologics	61,300	76,300	68,900	46,000	63,300
AE costs included	34,700	39,900	36,800	22,500	38,800
No offset costs	36,900	41,400	38,600	23,600	41,000
Extra cost for Pall	33,400	37,800	35,000	20,100	37,600
No negative QoL scores	48,600	56,500	52,100	30,700	52,800
Linear equation HAQ to QoL	38,600	43,800	40,600	23,700	42,300

Small variations in results where neither strategy had changed parameters reflect the first-and second-order sampling in the model.

TABLE 94 Results from scenario analysis: comparisons of other biologics against RTX (ICER in £/QALY)

Scenario	ADA–RTX	ETN–RTX	IFX–RTX	ABT–RTX
Reference	RTX	RTX	RTX	130,600
Vary time on TNF inhibitors	RTX	RTX	*4,100*	131,800
Same time on all biologics	206,000	RTX	RTX	131,200
RTX cycle time 6 months	*430*	RTX	*14,700*	51,500
RTX cycle time 11.6 months	RTX	RTX	RTX	861,100
Poor late DMARDs	RTX	RTX	RTX	158,600
HAQ change on biologics	RTX	RTX	RTX	96,400
AE costs included	RTX	RTX	RTX	126,100
No offset costs	RTX	RTX	RTX	134,100
Extra cost for Pall	RTX	RTX	RTX	131,000
No negative QoL scores	RTX	RTX	RTX	140,700
Linear equation HAQ to QoL	RTX	RTX	RTX	130,900

Incremental cost-effectiveness ratio in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective).

TABLE 95 Comparisons between biologics other than RTX (ICER in £/QALY)

Scenario	ADA–ABT	ETN–ABT	IFX–ABT	ADA–ETN	ADA–IFX	ETN–IFX
Reference	*46,400*	*37,800*	*41,700*	ADA	20,500	456,700
Vary time on TNF inhibitors	*47,700*	*38,900*	*39,100*	*72,800*	28,700	39,300
Same time on all biologics	*84,100*	*42,700*	*53,700*	ADA	21,600	351,500
RTX cycle time 6 months	*46,300*	*37,800*	*42,000*	ADA	21,700	1,325,400
RTX cycle time 11.6 months	*46,400*	*37,800*	*41,800*	ADA	20,700	591,000
Poor late DMARDs	*40,100*	*33,500*	*36,900*	ADA	20,600	316,000
HAQ change on biologics	*66,500*	*50,600*	*57,600*	ADA	24,300	IFX
AE costs included	*46,700*	*37,400*	*41,700*	ADA	19,000	502,600
No offset costs	*49,000*	*40,500*	*44,400*	ADA	23,500	460,000
Extra cost for Pall	*45,800*	*37,300*	*41,200*	ADA	20,300	452,000
No negative QoL scores	*60,300*	*48,300*	*53,700*	ADA	25,300	7,430,000
Linear equation HAQ to QoL	*49,100*	*40,300*	*44,600*	ADA	23,100	667,000

Incremental cost-effectiveness ratio in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). Small variations in results where neither strategy had changed parameters reflect the first-and second-order sampling in the model. It should be stressed that the comparisons between TNF inhibitors are based in each case (except 'Vary time on TNF inhibitors') on the mean values from a distribution, which covers all four quadrants of the cost-effectiveness plane.

with conventional DMARDs alone, the ICER for RTX is somewhat lower than for the other biologics. RTX dominates the TNF inhibitors (lower cost and more QALYs). The ICER for ABT compared with RTX is over £100,000/QALY. These results are subject to considerable uncertainty. Important drivers of that uncertainty were found in the scenario analysis to include:

- the assumptions about HAQ progression on biologic treatments
- the equation relating HAQ to QoL – in particular whether negative QoL scores can be allowed
- for comparisons involving RTX, the assumed time between treatments.

The results were fairly sensitive to the assumptions on efficacy of conventional DMARDs given after biologic therapy. The inclusion of AE costs for biologic therapy made little difference to the results. The mean time on RTX was considerably longer than for other biologics. This parameter was varied downwards in the scenario analysis 'Same time on all biologics' and the results were not generally sensitive to this parameter: this makes sense because the costs and QALYs in the RTX strategy were both reduced when the mean time on RTX was reduced.

Additional sensitivity analysis to assess impact of differences in assumptions between models

The main aim of this analysis was to explore the differences between the results of the various models. Two of the industry submissions (Abbott and Schering-Plough Ltd) contained ICERs that are directly comparable with the main BRAM results. Roche gave ICERs for RTX against DMARDs and against other biologics. As the mean costs and QALYs for RTX were the same in each comparison (*Tables 101–105* from MS, pp. 226–8 of their report), it is possible to infer the ICERs for other biologics against DMARDs. *Table 96* shows the results from the various models.

As well as the BRAM reference case, the scenario analysis with reduced efficacy for conventional DMARDs has been quoted above. This scenario is sufficient to account for cases where the Schering-Plough model gave a more favourable result than the BRAM reference case. Accordingly the main focus of further analysis is the assumptions in the Abbott and Roche models. Two aspects of the modelling have been considered: the short-term change in HAQ on starting treatment and the proportion of early quitters. The aim was to apply the industry assumptions to the BRAM. The process for doing this is described below.

Short-term change in Health Assessment Questionnaire on starting treatment

The Abbott and Roche models each had HAQ change based on ACR response using values shown in the tables below. To compare with the BRAM, it is necessary to convert this HAQ change pattern into a set of figures in the same structure as the BRAM. This means estimating *a* and *b* parameters for the beta distribution of HAQ change multipliers used in the BRAM. As the purpose of this exercise is to assess the impact of the difference in the effectiveness assumption, the mean HAQ change multiplier was estimated from the two company submissions. The value

TABLE 96 Comparison of model results

Model	ADA–DMARDs	ETN–DMARDs	IFX–DMARDs	RTX–DMARDs	ABT–DMARDs
BRAM reference	34,300	38,900	36,100	21,100	38,400
BRAM with poor late DMARDs	28,100	31,100	28,800	16,300	32,100
Abbott	16,000	16,000	21,500	11,000	30,100
Roche[a]	14,600	18,000	16,200	5,300	21,500
Schering-Plough Ltd	35,100	35,900	28,700	17,400	44,800

a Results for comparisons not involving RTX are inferred from total costs and QALYs reported by Roche (Tables 101–105 from MS, pp. 226–8 of the company submission).

TABLE 97 Probability of ACR responses based on the Abbott company submission

Treatment	ACR < 20	ACR20–50	ACR50–70	ACR > 70
TNF inhibitor	0.3574	0.2414	0.1958	0.2054
RTX	0.3822	0.2337	0.1858	0.1983
ABT	0.4531	0.2355	0.1631	0.1483
DMARDs	0.7474	0.1486	0.0631	0.0409

Calculated from figures in Table 3.2.3.1 of the Abbott submission (p. 51).

TABLE 98 Relative change in HAQ score by ACR response

Treatment	ACR < 20	ACR20–50	ACR50–70	ACR > 70
Biologics	0.110	0.405	0.588	0.806
DMARDs	0.016	0.300	0.565	0.735

Taken from Table 3.2.5.1 of the Abbott submission (p. 52).

of $a + b$ used in the BRAM reference case was preserved and the a and b parameters were inferred using this value.

For Abbott, the relevant figures were taken to be the ACR response rates (Table 3.2.3.1 of MS, p. 51) and the relative change in HAQ score based on ACR response by treatment from baseline to 6 months (Table 3.2.5.1 of MS, p. 52).

These are repeated for convenience (*Tables 97* and *98*).

The mean change in HAQ score for each type of treatment is then found from using the probabilities in *Table 97* as weights to calculate a weighted average of the changes in *Table 98*. For example, for TNF inhibitors, the calculation is $0.3574 \times 0.110 + 0.2414 \times 0.405 + 0.1958 \times 0.588 + 0.2054 \times 0.806 = 0.418$.

For ETN and IFX we have $a + b = 0.96$ from the reference case in the BRAM, from which $a = 0.418 \times 0.96 = 0.401$ and hence $b = 0.559$. Similar principles apply to the other DMARDs and the results are shown in *Table 99*.

Similarly, using the Roche parameters, *Table 100* shows the probability of responses. For HAQ change, Roche give absolute falls in HAQ. These have been converted in *Table 101* to relative changes by dividing by 2, which is the mean starting HAQ in the BRAM reference case. Then the same system of calculations gives the results in *Table 102*.

Changing the proportion of early quitters

Another potentially important difference between the models is the proportion of people withdrawing from the treatment early. The Abbott model reference case used failure to achieve ACR50 response as the criterion for early withdrawal. Therefore, 59.88% of those starting a TNF inhibitor would not continue beyond 6 months (Abbott submission, p. 51). In the BRAM reference case, the corresponding figure (for ADA) is just under 24%, made up of the short-term withdrawals at 13 weeks (9.9%) and the first 13 weeks of the long-term survival curve (15.4% of the remaining 90.1%). Similar remarks apply to all other drugs.

TABLE 99 Disease-modifying antirheumatic drug effectiveness for approximate equivalence to Abbott model

Treatment	Mean	*a* + *b*	*a*	*b*
ETN/IFX	0.418	0.96	0.401	0.559
ADA	0.418	1.24	0.518	0.722
RTX	0.406	0.95	0.385	0.565
ABT	0.361	1.18	0.426	0.754
LEF	0.122	1.22	0.149	1.071
GST	0.122	1.15	0.141	1.009
CyA	0.122	0.39	0.048	0.342
AZA	0.122	1.00	0.122	0.878

Means calculated as shown in *Tables 96* and *97*. Values of *a* + *b* preserved from BRAM reference case.

TABLE 100 Assumed ACR response rates in Roche model

Treatment	ACR < 20	ACR20–50	ACR50–70	ACR > 70
ADA	0.538	0.154	0.182	0.126
ETN	0.552	0.196	0.154	0.098
IFX	0.58	0.189	0.133	0.098
RTX	0.54	0.23	0.09	0.14
ABT	0.57	0.21	0.14	0.08
DMARDs	0.85	0.11	0.03	0.01

Source: Table 85 from Roche submission (p. 205).

TABLE 101 Relative change in HAQ score by ACR response in Roche model

ACR<20	ACR20–50	ACR50–70	ACR70+
0.05	0.225	0.405	0.555

Calculated from values in Table 86 of Roche submission (p. 206).

TABLE 102 Disease-modifying antirheumatic drug effectiveness for approximate equivalence to Roche model

Treatment	Mean	*a* + *b*	*a*	*b*
ADA	0.205	1.24	0.254	0.986
ETN	0.188	0.96	0.181	0.779
IFX	0.180	0.96	0.173	0.787
RTX	0.193	0.95	0.183	0.767
ABT	0.177	1.18	0.209	0.971
LEF	0.085	1.22	0.104	1.116
GST	0.085	1.15	0.098	1.052
CyA	0.085	0.39	0.033	0.357
AZA	0.085	1.00	0.085	0.915

Continuing to use ADA as the example, the BRAM reference case is based on a data set of 899 patients of whom 89 had withdrawn from treatment by 12 weeks, 50 of these for toxicity. For this exploratory analysis, the withdrawal time is changed to 26 weeks, and parameters for beta distributions are calculated on the basis of still having 899 patients of whom $0.5988 \times 899 = 538.3$ withdrew by 26 weeks. In the absence of any obvious alternative figure, the number withdrawing from because of toxicity is kept at 50. While rounding to the nearest integer would make little difference, the beta distributions can be used with non-integer parameters and so unrounded figures have been used. With regard to conventional DMARDs, the proportions withdrawing for toxicity either side of the 6-week cut-off have been maintained, the additional withdrawal rate being assigned to those withdrawal because of loss of effectiveness at 26 weeks.

As with the reference case, the structure of the model does not allow early withdrawal for RTX. *Table 103* shows the revised parameters.

For the Roche model, the early withdrawal rates were taken as the failure to achieve an ACR20 response and are therefore shown in *Table 100*. The same method was used to produce the figures in *Table 104*.

Results

For comparison with the Abbott model, the parameters in *Table 99* (short-term HAQ increase) were used in place of the BRAM reference case parameters in one analysis, all other parameters remaining as in the BRAM reference case. Separately, the parameters in *Table 103* (early withdrawal) were used, keeping the short-term HAQ increase as in the BRAM reference case. Finally, both sets of parameters were changed at the same time. The results are shown in *Table 105*. For comparison with the Roche model, the results of a similar analysis using the parameters in *Tables 102* and *104* are shown in *Table 106*.

TABLE 103 Early withdrawal for approximate equivalence to Abbott model

Treatment	Parameter	Point estimate (%)	Distribution
ADA	Withdrawal at 26 weeks	59.88	Beta (538.3, 360.7)
	Toxicity if above	9.29	Beta (50, 488.3)
ETN	Withdrawal at 26 weeks	59.88	Beta (243.1, 162.9)
	Toxicity if above	1.45	Beta (2, 136.2)
IFX	Withdrawal at 26 weeks	59.88	Beta (7.8, 5.2)
	Toxicity if above	25.61	Beta (2, 5.8)
RTX	No early withdrawal (see text)		
ABT	Withdrawal at 26 weeks	68.86	Beta (177.7, 80.3)
	Toxicity if above	5.08	Beta (9, 168.2)
LEF	Withdrawal at 6 weeks	13	Beta (13, 87)
	Withdrawal 6–26 weeks	76.6	Beta (76.6, 23.4)
	Toxicity if above	13.05	Beta (10, 66.6)
GST	Withdrawal at 6 weeks	14	Beta (10.1, 61.9)
	Withdrawal 6–26 weeks	75.6	Beta (54.4, 17.6)
	Toxicity if above	23.81	Beta (13, 41.6)
CyA	Withdrawal at 6 weeks	8	Beta (16, 184)
	Withdrawal 6–26 weeks	81.6	Beta (163.2, 36.8)
	Toxicity if above	14.71	Beta (24, 139.2)
AZA	Withdrawal at 6 weeks	15	Beta (15, 85)
	Withdrawal 6–26 weeks	74.6	Beta (74.6, 25.4)
	Toxicity if above	16.76	Beta (12.5, 62.1)

TABLE 104 Early quits for approximate equivalence to Roche model

Treatment	Parameter	Point estimate (%)	Distribution
ADA	Withdrawal at 26 weeks	53.8	Beta (483.7, 415.3)
	Toxicity if above	10.34	Beta (50, 433.6)
ETN	Withdrawal at 26 weeks	55.2	Beta (224.1, 181.9)
	Toxicity if above	1.57	Beta (2, 125.4)
IFX	Withdrawal at 26 weeks	58	Beta (7.5, 5.5)
	Toxicity if above	26.44	Beta (2, 5.6)
RTX	No early withdrawal (see text)		
ABT	Withdrawal at 26 weeks	57	Beta (147.1, 110.9)
	Toxicity if above	6.13	Beta (9, 137.7)
LEF	Withdrawal at 6 weeks	13	Beta (13, 87)
	Withdrawal 6–26 weeks	72	Beta (72, 28)
	Toxicity if above	13.89	Beta (10, 62)
GST	Withdrawal at 6 weeks	14	Beta (10.1, 61.9)
	Withdrawal 6–26 weeks	71	Beta (51.1, 20.9)
	Toxicity if above	25.35	Beta (13, 38.3)
CyA	Withdrawal at 6 weeks	8	Beta (16, 184)
	Withdrawal 6–26 weeks	77	Beta (154, 46)
	Toxicity if above	15.58	Beta (24, 130)
AZA	Withdrawal at 6 weeks	15	Beta (15, 85)
	Withdrawal 6–26 weeks	70	Beta (70, 30)
	Toxicity if above	17.86	Beta (12.5, 57.5)

TABLE 105 Comparison of model results with assumptions from Abbott model

Model	ADA–DMARDs	ETN–DMARDs	IFX–DMARDs	RTX–DMARDs	ABT–DMARDs
BRAM reference	34,300	38,900	36,100	21,100	38,400
Changing HAQ increase	21,700	21,900	20,100	11,100	28,700
Changing short-term withdrawal rate	22,200	23,400	26,200	19,500	24,100
Changing both	16,200	15,700	16,500	11,500	33,400
Abbott model	16,000	16,000	21,500	11,000	30,100

TABLE 106 Comparison of model results with assumptions from Roche model

Model	ADA–DMARDs	ETN–DMARDs	IFX–DMARDs	RTX–DMARDs	ABT–DMARDs
BRAM reference	34,300	38,900	36,100	21,100	38,400
Changing HAQ increase	31,900	33,500	32,000	17,200	41,500
Changing short-term withdrawal rate	23,800	25,100	27,400	20,500	26,400
Changing both	24,500	24,400	26,900	17,900	30,900
Roche[a]	14,600	18,000	16,200	5,300	21,500

a Results for comparisons not involving RTX are inferred from total costs and QALYs reported by Roche (*Tables 101–105*, p. 226–8 of the company submission).

Conclusion

The differences between the reference case results in the BRAM and those produced by Abbott and Schering-Plough Ltd can be explained by changing a small number of parameters in the model. There are some differences with the Roche model that remain unexplained in this analysis. It should be stressed that the purpose of this analysis is to compare the models and this is a separate matter from the discussion of the appropriateness of the various parameters.

Chapter 5

Assessment of factors relevant to the NHS and other parties

Wide use of biologic agents, NICE guidance on RA and the recent NAO report on services for patients with RA have profound implications for specialist rheumatology services. The NAO report suggests that acute trusts and primary care trusts (PCTs) have not yet met all the challenges they face. For example, monthly review in patients with active disease, as recommended in NICE guidance, is achieved by only 15% of acute trusts surveyed by the NAO. The main barriers reported by trusts were staffing, limited outpatient capacity and pressures to improve the ratio of follow-up to new patients. A majority of the acute trusts reported that they were unable to provide adequate follow-up for RA patients.[6] Models of shared care between primary care and secondary care exist, but only around half of the GPs in the NAO survey said that they had a shared care agreement with their local acute trust.[207] Good shared-care schemes with appropriate patient selection[71,208] could reduce the burden on specialists and meet some of the objectives set out in Lord Darzi's review.[209]

Increasing use of biologics, different mechanisms for obtaining funding (including appeals processes and inconsistency of response) for different PCTs and collection and submission of audit data have increased the administrative burden on specialist departments. PCTs have parallel demands with a need to monitor high-cost drug use and manage the implications of burgeoning NICE guidance while facing increasing demands from patients and hospital doctors with varying approaches to disease management. Expert teams remain vital to the delivery of services for RA patients, but pressures to provide community clinics in many locations risk fragmenting small teams and diluting expertise. The increasing complexity of care driven by new agents and more aggressive disease management means that primary care physicians are less able to take a lead role in the management of individual patients.[207] Also, the fact that prescriptions for biologics can be issued only by a specialist means that even better links between primary and secondary care colleagues are needed to co-ordinate care and avoid drug interactions.

Abatacept and TOC both require monthly i.v. infusions. Currently, such treatment is delivered largely in a hospital day-case unit. Capacity is under pressure as newer agents arrive and indications for existing agents widen. Solutions to improve capacity are needed. It seems likely that periodic i.v. infusions, required long term, will be administered away from acute hospitals and within patients' homes or other community settings. Pilot studies exploring IFX infusions at home in stable clients are under way.

In summary, it is imperative that acute trusts and PCTs are better placed to meet the challenges of therapeutic innovations in RA and the deficiencies of care identified by the NAO.

Chapter 6

Discussion

Statement of principal findings

Quantity and quality of evidence

Thirty-five studies described in 44 papers met the inclusion criteria. These included five RCTs, three comparative studies and 28 uncontrolled studies. Comparisons made in the included RCTs were switching to IFX (from ongoing ETN) versus ongoing ETN (OPPOSITE trial, $n = 27$);[133] RTX versus placebo with ongoing traditional DMARDs (REFLEX trial, $n = 517$);[124–126] ABT versus placebo with ongoing traditional DMARDs (ATTAIN trial, $n = 391$);[127–132] ABT added to ongoing ETN versus ongoing ETN (Weinblatt *et al.*,[134] $n = 121$);[40] and ABT added to ongoing biologics or non-biologic DMARDs versus ongoing biologics or non-biologic DMARDs (ASSURE trial, $n = 167$).[134] No directly relevant head-to-head trial directly comparing any of the five technologies against each other or directly comparing any of the technologies against other biologics or previously untried, newly initiated DMARDs, was found.

Effectiveness of adalimumab

No RCT was identified. Five uncontrolled studies with duration of follow-up ranging from 3 to 12 months showed that between 46% and 75% of patients achieved ACR20 and between 13% to 33% patients achieved ACR70. Mean reductions of 1.3–1.9 in DAS28 score and of 0.21–0.48 in HAQ score were observed. Results were not pooled owing to substantial clinical and statistical heterogeneity.

Effectiveness of etanercept

No RCT was found. Seven uncontrolled studies with duration of follow-up ranging from 3 to over 9 months showed that ACR20 was achieved in 37%–71% of patients after switching to ETN, ACR70 in 4%–21% of patients. Mean reductions of 0.47 to 1.80 in DAS28, and of 0.35 to 0.45 in HAQ score were observed. Results were not pooled due to substantial clinical and statistical heterogeneity between studies.

Effectiveness of infliximab

One RCT (OPPOSITE trial[133]) compared switching to IFX ($n = 13$) versus staying on ETN ($n = 14$) in patients who had an incomplete response to ETN. The study was considered not directly relevant to this report. Three uncontrolled studies with unclear length of follow-up were found, but none of these reported ACR response criteria or quantitative results of changes in DAS28 and HAQ scores.

Effectiveness of tumour necrosis factor inhibitors as a class

Some of the included studies assessed switching to an alternative TNF inhibitor, but did not provide data separately for individual TNF inhibitors. Two non-randomised comparative studies and six uncontrolled studies with duration of follow-up ranging from 3 months to 4 years were identified. ACR responses were reported in only one study, with response rates of 49% for ACR20 and 7% for ACR70 being observed. Reported mean reductions in DAS28 score ranged from −0.88 to −1.00. Only one study (using data from BSRBR) reported mean reduction in HAQ score of −0.11.

Effectiveness of rituximab

One good-quality RCT (REFLEX)[124–126] compared RTX with placebo (with ongoing DMARDs in both groups) in patients who had had inadequate response to one or more TNF inhibitors. At 6 months significantly more patients treated with RTX achieved ACR20 (RR = 2.85, 95% CI 2.08 to 2.91) and ACR70 (RR = 12.14, 95% CI 2.96 to 49.86) than those treated with the placebo. Significant differences between groups in favour of RTX were observed at 6 months for mean change from baseline in DAS28 score (mean difference –1.50, 95% CI –1.74 to –1.26) and mean change from baseline in HAQ score (mean difference –0.30, 95% CI –0.40 to –0.20). No significant difference in the risk of serious AEs and serious infections was observed. One non-randomised comparative study, five uncontrolled studies and two further analyses of data from RTX RCTs were also identified. Results generally supported findings from the REFLEX trial.[124–126]

Effectiveness of abatacept

One good-quality RCT (ATTAIN[127–132]) compared ABT with placebo (with ongoing DMARDs in both groups) in patients who had had inadequate response to one or more TNF inhibitors. At 6 months significantly more patients treated with ABT achieved ACR20 (RR = 2.56, 95% CI 1.77 to 3.69) and ACR70 (RR = 6.70, 95% CI 1.62 to 27.80) than those treated with the placebo. Significant differences between groups in favour of ABT were observed at 6 months for mean change from baseline in DAS28 score (mean difference –1.27, 95% CI –1.62 to –0.93) and mean change from baseline in HAQ score (mean difference –0.34, insufficient data for calculating 95% CI). No significant difference in the risk of serious AEs and serious infections was observed. Further data from the LTE of the ATTAIN trial[119] and a large prospective uncontrolled study (ARRIVE) generally supported findings from the ATTAIN trial.[127–132] Two further RCTs (Weinblatt *et al.*[133] and ASSURE[135]) were identified that compared ABT added to ongoing TNF inhibitors/biologics versus ongoing TNF inhibitors/biologics. The results from these trials showed patients who received a combination of ABT and a TNF inhibitor had an increased risk of infection and serious infection. This is reflected in the licensed indication, which advises against the use of such combination therapy, and thus further data from combination therapy were not assessed in this report.

Comparative effectiveness

No RCT provided evidence on genuine head-to-head comparisons between the technologies, other biologics and newly initiated, previously untried DMARDs. One non-randomised controlled study[136,137] compared switching to RTX versus switching to an alternative TNF inhibitor. The mean change in DAS28 score was greater in the RTX group than in the TNF inhibitor group (mean difference –0.35, 95% CI –0.71 to 0.01; median follow-up 11 months) but the difference just failed to reach statistical significance.

It was possible to carry out adjusted IC between RTX and ABT using data from placebo-controlled trials that included similar patient populations. The results showed no evidence of significant difference in their effectiveness (ACR20 for RTX vs ABT, RR = 1.12, 95% CI 0.68 to 1.84). No further analyses for comparative effectiveness were performed owing to limitation in available data.

Subgroup analyses

Evidence from the REFLEX trial[124–126] suggested that the effectiveness of RTX does not vary significantly according to reasons of withdrawal, baseline RF status and number of prior TNF inhibitors tried (one vs more than one).

No significant differences in the effectiveness of ABT between subgroups defined by the number of prior TNF inhibitor (one vs two) and the identity of the prior TNF inhibitor received (ETN vs

infliximab) were observed in the ATTAIN trial.[127–132] However, some of these subgroup analyses may be underpowered.

Evidence from observational studies showed that the proportion of patients responding to a subsequent TNF inhibitor might vary according to reason for withdrawal of the previous TNF inhibitor (higher response in patients who withdrew due to intolerance/AEs than in those who withdrew due to lack of efficacy). The proportion of patients who respond to a subsequent treatment (including TNF inhibitors, RTX and ABT) decreases as the number of prior TNF inhibitor(s) that the patients have tried increases.

Review of cost-effectiveness studies

Four studies met the inclusion criteria. All studies used a decision-analytic model. Published models vary in some important aspects: the type of model used, the sequence of drugs, comparator therapies and time horizon. All but one study carried out a cost–utility analysis and reported results in 'cost per QALY'. One study carried out a cost-effectiveness analysis and reported results in cost per additional case of 'low disease activity state' gained (DAS28 less than 2.6) and cost per additional remission gained (DAS28 less than or equal to 3.2). Appropriate sensitivity analyses were carried out in all studies. A comparison of ICERs between studies is not possible because of the different approaches to modelling, in particular time horizon, country of origin and perspective chosen. There was disparity in the selection of perspectives chosen for the analyses. One study reported costs that include both those from a health-care perspective as well as indirect costs and costs of informal care; inclusion of these costs improves the cost-effectiveness of the drug.

Independent modelling

The reference case model results show similar costs and QALYs for the TNF inhibitors, with somewhat lower costs and QALYs for RTX and higher costs and QALYs for ABT. Compared with conventional DMARDs alone, the ICER for RTX is somewhat lower than for the other biologics. RTX dominates the TNF inhibitors and the ICER for ABT compared with RTX is over £100,000/QALY. These results are subject to considerable uncertainty. Important drivers of that uncertainty were found in scenario analysis to include:

- the assumptions used about HAQ progression on biologic treatments
- the equation relating HAQ to QoL – in particular whether negative QoL scores can be allowed
- for comparisons involving RTX, the assumed time between treatments.

The inclusion of AE costs for biologic therapy made little difference to the results.

Strengths and limitations of the assessment

Strengths of the assessment

The strengths of this assessment include:

- A comprehensive literature review was undertaken which went beyond RCT evidence. Studies were selected and assessed according to a pre-specified protocol. Additional data from MSs were included.
- Key data were graphically presented in a systematic way to allow easy inspection of the variations between studies.
- Detailed subgroup analyses were carried out to examine factors that may influence the effectiveness of the technologies.

- The BRAM model has been further improved and modelling was carried out on various scenarios to explore uncertainties.

Limitation of the assessment

The limitations predominantly relate to factors outside the control of the assessment group. The major limitation of the assessment was the paucity of evidence from RCTs for assessing the clinical effectiveness of the three TNF inhibitors, and a complete absence of genuine head-to-head trials comparing the five technologies against each other, against other biologics or against newly initiated, previously untried DMARDs.

Given the paucity of RCT evidence, this report assessed data from observational studies that are more prone to potential bias. Most of the included studies were uncontrolled studies, which allow only the assessment of treatment response post-intervention compared with before intervention. Such comparisons do not adjust for the natural course of the disease; hence any observed responses could be attributed to possible effects of the treatment as well as other factors such as different methods of follow-up and data collection, data imputation and regression to the mean.

As registration of observational study is not mandated, they are more prone to publication bias. In addition, the reporting of outcomes varies widely between studies, and the scope for selective reporting of outcomes is substantial. These biases are difficult to assess.

The focus of this assessment was on the patient population who have had an inadequate response to a first TNF inhibitor. Many existing studies have included patient populations who withdrew from the previous TNF inhibitor due to AEs/intolerance and/or who had already tried more than one TNF inhibitor. The subgroup analysis suggests these factors may influence the proportion of patients who respond to subsequent treatments, but this does not necessarily translate into differential effectiveness measured as RR or RD. Furthermore, there is much less evidence to allow assessment of whether the magnitude of effects varies between subgroups in those patients who do respond. These require further research.

Uncertainties

Lack of good-quality evidence on effectiveness of the use of an alternative TNF inhibitor after patients had an inadequate response is the source of major uncertainty for this assessment. For the assessment of cost-effectiveness, lack of evidence assessing the effectiveness of previous untried traditional DMARDs in this patient population is also an important source of uncertainty.

Additional areas of uncertainty identified in the independent modelling include assumptions about HAQ progression on biologic treatments; whether negative QoL scores can be allowed when estimating QoL from HAQ score, and treatment interval between courses of RTX.

Chapter 7

Conclusions

Implications for service provision

In relation to the decision problems described in *Chapter 2*, the findings of this assessment report suggest:

1. There is a lack of good-quality evidence directly comparing the effectiveness of the five technologies against each other. This imposes significant uncertainties with regard to any assessment of their relative cost-effectiveness. Adjusted IC suggests that there is no significant difference in the effectiveness between RTX and ABT, both of which are supported by good-quality RCT evidence. Existing data do not allow reliable quantification of the effectiveness of TNF inhibitors compared with RTX and ABT. Independent modelling comparing each of the other four technologies with RTX (recommended in current NICE guidance) suggests RTX dominating ADA, ETN and infliximab, and an estimated ICER of £131,000 (per QALY) for ABT compared with RTX.
2. There is a lack of evidence comparing the effectiveness of the five technologies with newly initiated, previously untried DMARDs. Independent modelling based on certain assumptions suggests the following ICERs: £34,300 (per QALY) for ADA, £38,800 for ETN, £36,200 for infliximab, £21,200 for RTX and £38,600 for ABT.
3. There is a lack of evidence directly comparing the effectiveness of the five technologies with other biologic agents.
4. Good-quality evidence from RCTs suggests that RTX and ABT are more effective than supportive care (including ongoing DMARDs which had provided inadequate control of the disease). Data from observational studies suggest that the use of an alternative TNF inhibitor after patients had inadequate response to a first TNF inhibitor may offer some benefit, but there remain significant uncertainties with regard to the magnitude of treatment effects and how these translate into cost-effectiveness.
5. Good-quality evidence from RCTs does not suggest differential effectiveness between various subgroups for RTX and ABT.

Suggested research priorities

The following research priorities are suggested in view of findings of this assessment:

- Head-to-head trials of adequate size and duration comparing the clinical effectiveness and cost-effectiveness of the technologies against each other and emerging biologics.
- Good-quality studies collecting information on the clinical effectiveness and cost-effectiveness of the technologies compared with previously untried conventional DMARDs in this patient population.
- Further analysis and synthesis of existing and future RCT data to quantify the potential impact of reasons for withdrawal of first TNF inhibitor, the history of prior exposure to TNF inhibitor(s) and autoantibody status (e.g. RF and anti-CCP antibody) on the effectiveness of the technologies.

- An overarching synthesis of evidence for the effectiveness of treatment modalities that can be used in various places of the treatment pathway for RA.
- Development of technologies/methods for identifying patients who are likely to respond to a biologic with a particular mode of action.
- Assessment of different methods and tariffs of utility valuations in RA and the impact of different methods on economic evaluation.

Acknowledgements

We thank Karen Biddle for providing administrative support.

Contributions of authors

Kinga Malottki was the main reviewer on this report and maintained day-to-day running of the review. She participated in study selection, data extraction and analyses. She drafted the methods results (template) and sections and edited the report. She conducted the clinical analyses for infliximab, abatacept and comparative studies.

Dr Pelham Barton constructed and revised the Birmingham Rheumatoid Arthritis Model (BRAM) and carried out de novo modelling using the revised BRAM. He wrote the sections of the report relating to modelling and also provided senior support for all economic sections.

Angelos Tsourapas conducted the cost-effectiveness review and critique of MSs.

Abdulrahman Uthman participated in data extraction and data checking and conducted analyses for etanercept and TNF inhibitors as a class.

Zulian Liu participated in data extraction and data checking and conducted analyses for adalimumab and rituximab.

Dr Kristina Routh participated in data extraction and data checking.

Dr Martin Connock provided support for statistical analyses and conducted indirect analyses.

Dr Paresh Jobanputra provided clinical advice, conducted the survey of the West Midlands Rheumatologists and drafted the background and factors relevant to the NHS sections.

Dr David Moore participated in study selection, edited various sections of the report and provided senior support.

Anne Fry-Smith devised and implemented search strategies for bibliographic databases and drafted the searching methods section.

Dr Yen-Fu Chen was the senior reviewer on this report and provided project management and advice on all aspects of the report. He compiled the study protocol, participated in study selection, data extraction and analyses, conducted subgroup analyses, drafted the summary and discussion and takes responsibility for the whole report.

References

1. Jacoby R, Jayson M, Cosh J. Onset, early stages, and prognosis of rheumatoid arthritis: a clinical study of 100 patients with 11-year follow-up. *Br Med J* 1973;**ii**:96–100.
2. Wiles NJ, Scott DGI, Barrett EM, Merry P, Arie E, Gaffney K, *et al.* Benchmarking: the five year outcome of rheumatoid arthritis assessed using a pain score, the Health Assessment Questionnaire and the Short Form-36 in a community and a clinic based sample. *Ann Rheum Dis* 2001;**60**:956–61.
3. Arnett FC, Edworthy S, Bloch D. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315–24.
4. Wiles N, Symmons D, Harrison B, Barrett E, Barrett J, Scott D. Estimating the incidence of rheumatoid arthritis. Trying to hit a moving target? *Arthritis Rheum* 1999;**42**:1339–46.
5. 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.
6. National Audit Offices. *Services for people with rheumatoid arthritis*. The Stationary Office. URL: www.nao.org.uk/publications/0809/services_for_people_with_rheum.aspx (accessed November 2009).
7. MacGregor A, Snieder H, Rigby A, Koskenvuo M, Kapiro J, Aho K. Characterising the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum* 2000;**43**:30–7.
8. Jones MA, Silman AJ, Whiting S, Barrett EM, Symmons DP. Occurrence of rheumatoid arthritis is not increased in the first degree of relatives of a population based inception cohort of inflammatory polyarthritis. *Ann Rheum Dis* 1996;**55**:89–93.
9. Barton P, Worthington J. Genetic susceptibility to rheumatoid arthritis: an emerging picture. *Arthritis Rheum* 2009;**61**:1441–6.
10. Newton JL, Harvey SM, Wordsworth BP, Brown MA. A review of the MHC genetics of rheumatoid arthritis. *Genes Immun* 2004;**5**:151–7.
11. Brennan F, McInnes I. Evidence that cytokines play a role in rheumatoid arthritis. *J Clin Invest* 2008;**118**:3537–45.
12. Firestein GS. Pathogensis of rheumatoid arthritis: how early is early? *Arthritis Res Ther* 2005;**7**:157–9.
13. Edwards JCW, Cambridge G. Prospects for B-cell-targeted therapy in autoimmune disease. *Rheumatology* 2005;**44**:151–6.
14. Waldburger JM, Firestein GS. Garden of therapeutic delights: new targets in rheumatic diseases. *Arthritis Res Ther* 2009;**11**:206.
15. National Collaborating Centre for Chronic Conditions. *Rheumatoid arthritis: national clinical guideline for management and treatment in adults*. London: Royal College of Physicians; 2009.
16. Kirwan JR, Power LL. *Glucocorticoids in rheumatic disease*. URL: www.arc.org.uk/arthinfo/medpubs/6633/6633.asp (accessed November 2009).
17. Jones G, Halbert J, Crotty M, Shanahan EM, Batterham M, Ahern M. The effect of treatment on radiological progression in rheumatoid arthritis: a systematic review of randomized placebo-controlled trials. *Rheumatology* 2003;**42**:6–13.

18. Jobanputra P, Wilson J, Douglas K, Burls A. A survey of British rheumatologists' DMARD preferences for rheumatoid arthritis. *Rheumatology* 2004;**43**:206–10.

19. Pincus T, Marcum SB, Callahan LF. Long-term drug therapy for rheumatoid arthritis in seven rheumatology private practices: II. Second line drugs and prednisone. *J Rheumatol* 1992;**19**:1894.

20. Klarenbeek NR, Güler-Yüksel M, Gerards AH, Kerstens PJSM, Molenaar ETH, Huizinga TW, *et al.* Clinical and radiological outcomes of four DAS driven treatment strategies: 6-year results of the BeSt study. American College for Rheumatology, Annual Meeting, Philadephia, PA, 2009; abstract no. 1019.

21. Grigir C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, *et al.* Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;**364**:263–9.

22. Goekoop-Ruiterman YPM, De Vries-Bouwstra JK, Allaart CF, Van ZD, Kerstens PJSM, Hazes JMW, *et al.* Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2007;**146**:406–15.

23. Moreland LW, O'Dell JR, Paulus H, Curtis JR, Bridges SL, Jr, Zhang X, *et al.* Treatment of early aggressive Ra; a randomized, double-blind, 2-year trial comparing immediate triple DMARD versus MTX plus etanercept to step-up from initial MTX monotherapy. American College for Rheumatology, Annual Meeting, Philadephia, PA, 2009; abstract no. 1895.

24. Wiles N, Scott DGI, Barrett EM, Merry P, Arie E, Gaffney K, *et al.* Benchmarking: the five year outcome of rheumatoid arthritis assessed using a pain score, the Health Assessment Questionnaire, and the Short Form-36 (SF-36) in a community and a clinic based sample. *Ann Rheum Dis* 2001;**60**:956–61.

25. 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.

26. Smolen JS, Aletaha D, Keystone E. Superior efficacy of combination therapy for rheumatoid arthritis fact or fiction? *Arthritis Rheum* 2005;**52**:2975–83.

27. Saag KG, Gim GT, Patkar NM, Anuntiyo J, Finney C, Curtis JR, *et al.* American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2008;**59**:762–84.

28. 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.

29. National Institute for Health and Clinical Excellence. *Rheumatoid arthritis: consultation table.* URL: www.nice.org.uk/guidance/index.jsp?action=download&o=43340 (accessed November 2009).

30. Treharne GJ, Lyons AC, Hale ED, Douglas KMJ, Kitas GD. 'Compliance' is futile but is 'concordance' between rheumatology patients and health professionals attainable? *Rheumatol* 2006;**45**:1–5.

31. Scott D, Shipley M, Dawson A, Edwards S, Symmons D, Woolf A. The clinical management of rheumatoid arthritis and osteoarthritis: strategies for improving clinical effectiveness. *Br J Rheumatol* 1998;**37**:546–54.

32. Scottish Intercollegiate Guidelines Network (SIGN). *Management of early rheumatoid arthritis*. URL: www.sign.ac.uk/guidelines/fulltext/48/index.html (accessed 17 May 2005).

33. Wolfe F, Zwillich SH. The long-term outcomes of rheumatoid arthritis. *Arthritis Rheum* 1998;**41**:1072–82.

34. van Gestel AM, Haagsma CJ, Van Riel PLCM. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;**41**:1845–50.

35. Fransen J, Creemers MCW, Van Riel PLCM. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology* 2004;**43**:1252–55.

36. Jerram S, Butt S, Gadsby K, Deighton C. Discrepancies between the EULAR response criteria and the NICE guidelines for continuation of anti-TNF therapy in RA: a cause for concern? *Rheumatology* 2008;**47**:180–2.

37. Wolfe F, Michaud K, Pincus T, Furst D, Keystone E. The disease activity score is not suitable as the sole criterion for initiation and evaluation of anti-tumor necrosis factor therapy in the clinic. *Arthritis Rheum* 2005;**52**:3873–9.

38. Uhlig T, Kvien TK, Pincus T. Test retest reliability of disease activity core set measures and indicies in rheumatoid arthritis. *Ann Rheum Dis* 2009;**68**:972–5.

39. Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20). *J Rheumatol* 2001;**28**:982–9.

40. Kirwan JR. Links between radiological change, disability and pathology in rheumatoid arthritis. *Rheumatology* 2001;**28**:881–6.

41. Landewe R, Boers M, van der Heijde D. How to interpret radiological progression in randomised clinical trials? *Rheumatology* 2003;**42**:2–5.

42. Young A, Dixey J, Cox N, Davies P, Devlin J, Emery P. How does disability in early rheumatoid arthritis (RA) affect patients and their lives? Results of 5 years follow-up in 732 patients from the Early RA Study (ERAS). *Rheumatology* 2000;**39**:603–11.

43. Barrett EM, Scott DGI, Wiles N, Symmons D. The impact of rheumatoid arthritis on employment status in the early years of disease: a UK community-based study. *Rheumatology* 2000;**39**:1403–9.

44. Chorus AMJ, Miedem HS, Wevers CJ, van der Linden S. Labour force participation among patients with rheumatoid arthritis. *Ann Rheum Dis* 2000;**59**:549–54.

45. Goodson N, Wiles N, Lunt M, Barrett EM, Silman A, Symmons DPM. Mortality in early inflammatory polyarthritis. Cardiovascular disease is increased in seropositive patients. *Arthritis Rheum* 2002;**46**:2010–19.

46. Boers M, Dijkmans B, Gabriel S, Maradit-Kremers H, O'Dell J, Pincus T. Making an impact on mortality in rheumatoid arthritis. Targeting cardiovascular disease. *Arthritis Rheum* 2004;**50**:1734–9.

47. Navarro-Cano G, del Rincon I, Pogosian S, Roldan JF, Escalante A. Association of mortality with disease severity in rheumatoid arthritis, independent of comorbidity. *Arthritis Rheum* 2003;**48**:2425–33.

48. Ward MM. Recent improvements in survival in patients with rheumatoid arthritis: better outcomes or different study designs? *Arthritis Rheum* 2001;**44**:1467–9.

49. Gabriel SE, Crowson CS, Kremers HM, Doran MF, Turesson C, O'Fallon WM, *et al.* Survival in rheumatoid arthritis: a population-based analysis of trends over 40 years. *Arthritis Rheum* 2003;**48**:54–8.

50. Menue C, Touze E, Ludovic 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.

51. Gabriel S. Why do people with rheumatoid arthritis still die prematurely? *Ann Rheum Dis* 2008;**67**:iii30–iii34.

52. Doran MF, Crowson CS, Pond GR, O'Fallon WM, Gabriel SE. Predictors of infection in rheumatoid arthritis. *Arthritis Rheum* 2002;**46**:2294–300.

53. Hallert E, Husberg M, Jonsson D, Skogh T. Rheumatoid arthritis is already expensive during the first year of the disease (the Swedish TIRA project). *Rheumatology* 2004;**43**:1374–82.

54. Brouwer WB, van Exel NJ, Van de Berg B, Dinant HJ, Koopmanschap MA. Burden of caregiving: evidence of objective burden, subjective burden, and quality of life impacts on informal caregivers of patients with rheumatoid arthritis. *Arthritis Rheum* 2004;**51**:570–7.

55. Michaud K, Messer J, Choi HK, Wolfe F. Direct medical costs and their predictors in patients with rheumatoid arthritis: a three-year study of 7,257 patients. *Arthritis Rheum* 2003;**48**:2750–62.

56. Puolakka K, Kautiainen H, Mottonen T, Hannonen P, Korpela M, Julkunen H, *et al.* Impact of initial aggressive drug treatment with a combination of disease-modifying antirheumatic drugs on the development of work disability in early rheumatoid arthritis: a five-year randomized follow-up trial. *Arthritis Rheum* 2004;**50**:55–62.

57. Young A, Dixey J, Kulinskay E, Cox N, Davies P, Devlin J, *et al.* Which patients stop working because of rheumatoid arthritis? Results of five years' follow up in 732 patients from Early RA Study (ERAS). *Ann Rheum Dis* 2002;**6**:335–40.

58. National Institute for Health and Clinical Excellence. *Certolizumab pegol for the treatment of rheumatoid arthritis*. NICE technology appraisal guidance 186. 2010. URL: http://guidance.nice.org.uk/TA186/Guidance/pdf/English (accessed 15 February 2011).

59. Ehlers S. Why does tumour necrosis factor targeted therapy reactivate tuberculosis? *J Rheumatol* 2005;**32**:35–9.

60. Kindler V, Sappiro AP, Grau GE, Piguet PF, Vassalli P. The inducing role of tumor necrosis factor in the development of bactericidal granulomas during BCG infection. *Cell* 1989;**56**:731–40.

61. British Thoracic Society Standards of Care Committee. BTS recommendations for assessing risk, and for managing M.tuberculosis infection and disease in patients due to start anti-TNF-alpha treatment. *Thorax* 2005;**60**:800–5.

62. Ledingham J, Wilkinson C, Deighton C. British Thoracic Society (BTS) recommendations for assessing risk and managing tuberculosis in patients due to start anti-TNF-a treatments. *Rheumatology* 2005;**44**:1205–6.

63. Gardam MA, Keystone EC, Menzies R, Manners S, Skamene E, Long R, *et al.* Anti-tumour necrosis factor agents and tuberculosis risk: mechanisms of action and clinical management. *Lancet* 2003;**3**:148–55.

64. Charles PJ, Smeenk RJT, De Jong J, Feldmann M, Maini R. Assessment of antibodies to double-stranded DNA induced in rheumatoid arthritis patients following treatment with

infliximab, a monoclonal antibody to tumour necrosis factor a: findings in open-label and randomised placebo controlled trials. *Arthritis Rheum* 2000;**43**:2383–90.

65. Fleischmann RM. Progressive multifocal leukoencephalopathy following rituximab treatment in a patient with rheumatoid arthritis. *Arthritis Rheum* 2009;**60**:3225–8.
66. National Institute for Health and Clinical Excellence. *Tocilizumab for the treatment of rheumatoid arthritis*. NICE technology appraisal guidance 198. 2010. URL: http://guidance.nice.org.uk/TA198/Guidance/pdf/English (accessed 15 February 2011).
67. Hetland ML, Stengaard-Pedersen K, Junker P, Lottenburger T, Hansen I, Andersen LS, *et al.* Aggressive combination therapy with intra-articular glucocorticoid injections and conventional disease-modifying anti-rheumatic drugs in early rheumatoid arthritis: second-year clinical and radiographic results from the CIMESTRA study. *Ann Rheum Dis* 2008;**67**:815–22.
68. Van Vollenhoven RF, Ernestam S, Geborek P, Petersson IF, Coster L, Waltbrand E, *et al.* Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in patients with early rheumatoid arthritis (Swefot trial): 1-year results of a randomised trial. *Lancet* 2009;**374**:459–66.
69. Emery P, Breedveld FC, Hall S, Durez P, Chang DJ, Robertson D, *et al.* Comparison of methotrexate monotherapy with a combination of methotrexate and etanercept in active, early, moderate to severe rheumatoid arthritis (COMET): a randomised, double-blind, parallel treatment trial. *Lancet* 2008;**372**:375–82.
70. Fraenkel L, Bogardus ST, Concato J, Felson DT, Wittink DR. Patient preferences for treatment of rheumatoid arthritis. *Ann Rheum Dis* 2004;**63**(11):1372–8.
71. Symmons D, Tricker K, Roberts C, Davies L, Dawes P, Scott DL. The British Rheumatoid Outcome Study Group (BROSG) randomised controlled trial to compare the effectiveness and cost-effectiveness of aggressive versus symptomatic therapy in established rheumatoid arthritis. *Health Technol Assess* 2005;**9**(34).
72. The British Society for Rheumatology Biologics Register. *Newsletter*. URL: www.medicine.manchester.ac.uk/images/File/bsrbr_newsletter_oct_2009.pdf (accessed November 2009).
73. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;**50**:683–91.
74. Suarez-Almazor M, Ortiz Z, Lopez-Olivo M, Pak C, Skidmore B, Kimmel B, *et al.* Infliximab and etanercept in rheumatoid arthritis: timing, dose escalation, and switching. *Ottawa: Canadian Agency for Drugs and Technologies in Health* (CADTH) 2007;**47**:1–34.
75. Anonymous. Rituximab (in rheumatoid arthritis): for a few patients, with close monitoring. *Prescrire Int* 2007;**16**:186–8.
76. Carmona L, Ortiz A, Abad MA. How good is to switch between biologics? A systematic review of the literature. *Acta Reumatol Port* 2007;**32**:113–28.
77. Unit of Health Economics and Technology Assessment. *Rituximab in patients with rheumatoid arthritis: systematic review and economic evaluation*. Budapest: Unit of Health Economics and Technology Assessment in Health Care (HUNHTA); 2006.
78. Kaine JL. Abatacept for the treatment of rheumatoid arthritis: a review. *Curr Ther Res Clin Exp* 2007;**68**:379–99.
79. Brodszky V, Czirjak L, Geher P, Hodinka L, Karpati K, Pentek M, *et al.* Rituximab in patients with rheumatoid arthritis: systematic review. *Orvosi Hetilap* 2007;**148**(40):1883–93.

80. Sherrer Y. Abatacept in biologic-naive patients and TNF inadequate responders: clinical data in focus. *Curr Med Res Opin* 2008;**24**:2283–94.

81. Lutt JR, Deodhar A. Rheumatoid arthritis: strategies in the management of patients showing an inadequate response to TNFalpha antagonists. *Drugs* 2008;**68**:591–606.

82. Coughlin M. Improving patient outlook in rheumatoid arthritis: experience with abatacept. *J Am Acad Nurse Pract* 2008;**20**:486–95.

83. Calvet FJ, Maymo J. What can be done in inadequate response to anti-TNF inhibitors? Review of the evidence. *Reumatologia Clinica Suplementos* 2008;**3**:24–31.

84. De Lara MGM, Balsa A. Efficacy of rituximab in rheumatoid arthritis. *Medicina Clinica Monografias* 2008;**9**:20–6.

85. Fernandez-Lopez C, Blanco FJ. ATTAIN study: efficacy of abatacept in patients with rheumatoid arthritis and inadequate response to anti-TNF-alpha. *Reumatologia Clinica Suplementos* 2006;**1**:34–43.

86. Maymo-Guarch J. Rituximab in patients with rheumatoid arthritis and inadequate response to anti-TNF therapy. *Reumatologia Clinica Suplementos* 2006;**1**:31–5.

87. Morovic-Vergles J. [Rituximab (Mabthera)--treatment of rheumatoid arthritis patients with inadequate response to TNF inhibitors--when to change therapy?]. [Croatian]. *Reumatizam* 2008;**55**:70–2.

88. Kristensen LE, Saxne T, Geborek P. Switching between anti-tnf therapies does not affect level of adherence to therapy in rheumatoid arthritis but response rates seem to decline. *Ann Rheum Dis* 2006;**65**:327.

89. Kafka SP, Hinkle K, Reed G. Discontinuing or switching TNF antagonists in patients with rheumatoid arthritis: data collected from the corrona database. *Ann Rheum Dis* 2005;**64**:467.

90. Naumann J, Detert J, Buttgereit F, Burmester G. A second anti-tnfa therapy after treatment failure of the first anti-tnfa therapy results in a significant decrease of concomitant glucocorticoid treatment in patients with rheumatoid arthritis. *Ann Rheum Dis* 2006;**65**:632.

91. Genentech. *A study of retreatment with rituximab in patients with rheumatoid arthritis receiving background methotrexate (SUNRISE)*. URL: http://clinicaltrials.gov/ct2/show/NCT00266227 (accessed July 2009).

92. Bennett AN, Peterson P, Zain A, Grumley J, Panayi G, Kirkham B. Adalimumab in clinical practice. Outcome in 70 rheumatoid arthritis patients, including comparison of patients with and without previous anti-TNF exposure. *Rheumatology* 2005;**44**:1026–31.

93. Wick MC, Ernestam S, Lindblad S, Bratt J, Klareskog L, Van Vollenhoven RF. Adalimumab (Humira) restores clinical response in patients with secondary loss of efficacy from infliximab (Remicade) or etanercept (Enbrel): results from the STURE registry at Karolinska University Hospital. *Scand J Rheumatol* 2005;**34**:353–8.

94. Nikas SN, Voulgari PV, Alamanos Y, Papadopoulos CG, Venetsanopoulou AI, Georgiadis AN, *et al.* Efficacy and safety of switching from infliximab to adalimumab: a comparative controlled study. *Ann Rheum Dis* 2006;**65**:257–60.

95. Burmester GR, Mariette X, Montecucco C, Monteagudo-Saez I, Malaise M, Tzioufas AG, *et al.* Adalimumab alone and in combination with disease-modifying antirheumatic drugs for the treatment of rheumatoid arthritis in clinical practice: the Research in Active Rheumatoid Arthritis (ReAct) trial. *Ann Rheum Dis* 2007;**66**:732–9.

96. Bombardieri S, Ruiz AA, Fardellone P, Geusens P, McKenna F, Unnebrink K, *et al.* Effectiveness of adalimumab for rheumatoid arthritis in patients with a history of TNF-antagonist therapy in clinical practice. *Rheumatology* 2007;**46**:1191–9.

97. van der Bijl AE, Breedveld FC, Antoni CE, Kalden JR, Kary S, Burmester GR, *et al.* An open-label pilot study of the effectiveness of adalimumab in patients with rheumatoid arthritis and previous infliximab treatment: relationship to reasons for failure and anti-infliximab antibody status. *Clin Rheumatol* 2008;**27**:1021–8.

98. Haraoui B, Keystone EC, Thorne JC, Pope JE, Chen I, Asare CG, *et al.* Clinical outcomes of patients with rheumatoid arthritis after switching from infliximab to etanercept. *J Rheumatol* 2004;**31**:2356–9.

99. Buch MH, Seto Y, Bingham SJ, Bejarano V, Bryer D, White J, *et al.* C-reactive protein as a predictor of infliximab treatment outcome in patients with rheumatoid arthritis: defining subtypes of nonresponse and subsequent response to etanercept. *Arthritis Rheum* 2005;**52**:42–8.

100. Cohen G, Courvoisier N, Cohen JD, Zaltni S, Sany J, Combe B. The efficiency of switching from infliximab to etanercept and vice-versa in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2005;**23**:795–800.

101. Buch MH, Bingham SJ, Bejarano V, Bryer D, White J, Emery P, *et al.* Therapy of patients with rheumatoid arthritis: outcome of infliximab failures switched to etanercept. *Arthritis Care Res (Hoboken)* 2007;**57**:448–53.

102. Iannone F, Trotta F, Monteccuco C, Giacomelli R, Galeazzi M, Matucci-Cerinic M. Erratum: Etanercept maintains the clinical benefit achieved by infliximab in patients with rheumatoid arthritis who discontinued infliximab because of side effects (*Ann Rheum Dis* 2007;**66**:249–52). *Ann Rheum Dis* 2007;**66**:847.

103. Laas K, Peltomaa R, Kautiainen H, Leirisalo-Repo M. Clinical impact of switching from infliximab to etanercept in patients with rheumatoid arthritis. *Clin Rheumatol* 2008;**27**:927–32.

104. Bingham CO, III, Ince A, Haraoui B, Keystone EC, Chon Y, Baumgartner S. Effectiveness and safety of etanercept in subjects with RA who have failed infliximab therapy: 16-week, open-label, observational study. *Curr Med Res Opin* 2009;**25**:1131–42.

105. Ang HTS, Helfgott S. Do the clinical responses and complications following etanercept or infliximab therapy predict similar outcomes with the other tumor necrosis factor-alpha antagonists in patients with rheumatoid arthritis? *J Rheumatol* 2003;**30**:2315–18.

106. Hansen KE, Hildebrand JP, Genovese MC, Cush JJ, Patel S, Cooley DA, *et al.* The efficacy of switching from etanercept to infliximab in patients with rheumatoid arthritis. *J Rheumatol* 2004;**31**:1098–102.

107. Yazici Y, Erkan D, Van Vollenhoven RF. Do etanercept-naive patients with rheumatoid arthritis respond better to infliximab than patients for whom etanercept has failed? *Ann Rheum Dis* 2004;**63**:607–8.

108. Gomez-Reino JJ, Carmona L, Biobadaser group. Switching TNF antagonists in patients with chronic arthritis: an observational study of 488 patients over a four-year period. *Arthritis Res Ther* 2006;**8**:R29.

109. Solau-Gervais E, Laxenaire N, Cortet B, Dubucquoi S, Duquesnoy B, Flipo R-M. Lack of efficacy of a third tumour necrosis factor alpha antagonist after failure of a soluble receptor and a monoclonal antibody. *Rheumatology* 2006;**45**:1121–4.

110. Hjardem E, Ostergaard M, Podenphant J, Tarp U, Andersen LS, Bing J, *et al.* Do rheumatoid arthritis patients in clinical practice benefit from switching from infliximab to a second tumor necrosis factor alpha inhibitor? *Ann Rheum Dis* 2007;**66**:1184–9.

111. Duftner C, Dejaco C, Larcher H, Schirmer M, Herold M. Biologicals in rheumatology: Austrian experiences from a rheumatic outpatient clinic. *Rheumatol Int* 2008;**29**:69–73.

112. Karlsson JA, Kristensen LE, Kapetanovic MC, Gulfe A, Saxne T, Geborek P. Treatment response to a second or third TNF-inhibitor in RA: results from the South Swedish Arthritis Treatment Group Register. *Rheumatology* 2008;**47**:507–13.

113. Blom M, Kievit W, Fransen J, Kuper IH, Den Broeder AA, De Gendt CM, *et al.* The reason of discontinuation of the first tumor necrosis factor (TNF) blocking agent does not influence effectiveness of a second TNF blocking agent in patients with rheumatoid arthritis. *J Rheumatol* 2009;**36**:2171–7.

114. Bokarewa M, Lindholm C, Zendjanchi K, Nadali M, Tarkowski A. Efficacy of anti-CD20 treatment in patients with rheumatoid arthritis resistant to a combination of methotrexate/anti-TNF therapy. *Immunol* 2007;**66**:476–83.

115. Jois RN, Masding A, Somerville M, Gaffney K, Scott DGI. Rituximab therapy in patients with resistant rheumatoid arthritis: real-life experience. *Rheumatology* 2007;**46**:980–2.

116. Keystone E, Fleischmann R, Emery P, Furst DE, van Vollenhoven R, Bathon J, *et al.* Safety and efficacy of additional courses of rituximab in patients with active rheumatoid arthritis: an open-label extension analysis. *Arthritis Rheum* 2007;**56**:3896–908.

117. Assous N, Gossec L, Dieude P, Meyer O, Dougados M, Kahan A, *et al.* Rituximab therapy in rheumatoid arthritis in daily practice. *J Rheumatol* 2008;**35**:31–4.

118. Thurlings RM, Vos K, Gerlag DM, Tak PP. Disease activity-guided rituximab therapy in rheumatoid arthritis: the effects of re-treatment in initial nonresponders versus initial responders. *Arthritis Rheum* 2008;**58**:3657–64.

119. Genovese MC, Schiff M, Luggen M, Becker J-C, Aranda R, Teng J, *et al.* Efficacy and safety of the selective costimulation modulator abatacept following 2 years of treatment in patients with rheumatoid arthritis and an inadequate response to anti-tumour necrosis factor therapy. *Ann Rheum Dis* 2008;**67**:547–54.

120. Schiff M, Pritchard C, Huffstutter JE, Rodriguez-Valverde V, Durez P, Zhou X, *et al.* The 6-month safety and efficacy of abatacept in patients with rheumatoid arthritis who underwent a washout after anti-tumour necrosis factor therapy or were directly switched to abatacept: the ARRIVE trial. *Ann Rheum Dis* 2009;**68**:1708–14.

121. Hyrich KL, Lunt M, Watson KD, Symmons DPM, Silman AJ. Outcomes after switching from one anti-tumor necrosis factor alpha agent to a second anti-tumor necrosis factor alpha agent in patients with rheumatoid arthritis: results from a large UK national cohort study. *Arthritis Rheum* 2007;**56**:13–20.

122. Hyrich KI, Lunt M, Dixon WG, Watson KD, Symmons DPM. Effects of switching between anti-TNF therapies on HAQ response in patients who do not respond to their first anti-TNF drug. *Rheumatology* 2008;**47**:1000–5.

123. NICE Decision Support Unit. *Effect of a second course of anti-TNF therapy on HAQ following lack of response to the first course*. URL: www.nice.org.uk/nicemedia/live/11702/37296/37296.pdf (accessed January 2010).

124. Keystone E, Burmester GR, Furie R, Loveless JE, Emery P, Kremer J, *et al.* Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid

arthritis refractory to anti-tumor necrosis factor therapy. *Arthritis Care Res (Hoboken)* 2008;**59**:785–93.

125. Keystone E, Emery P, Peterfy CG, Tak PP, Cohen S, Genovese MC, *et al.* Rituximab inhibits structural joint damage in patients with rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitor therapies. *Ann Rheum Dis* 2009;**68**:216–21.

126. Cohen SB, Emery P, Greenwald MW, Dougados M, Furie RA, Genovese MC, *et al.* Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. *Arthritis Rheum* 2006;**54**:2793–806.

127. Hassett AL, Li T, Buyske S, Savage SV, Gignac MAM. The multi-faceted assessment of independence in patients with rheumatoid arthritis: preliminary validation from the ATTAIN study. *Curr Med Res Opin* 2008;**24**:1443–53.

128. Emery P. Abatacept has beneficial effects in rheumatoid arthritis patients with an inadequate response to anti-TNFalpha therapy. *Clin Exp Rheumatol* 2005;**23**:767–8.

129. Westhovens R, Cole JC, Li T, Martin M, Maclean R, Lin P, *et al.* Improved health-related quality of life for rheumatoid arthritis patients treated with abatacept who have inadequate response to anti-TNF therapy in a double-blind, placebo-controlled, multicentre randomized clinical trial. *Rheumatology* 2006;**45**:1238–46.

130. Genovese MC, Becker J-C, Schiff M, Luggen M, Sherrer Y, Kremer J, *et al.* Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. *N Engl J Med* 2005;**353**:1114–23.

131. Bristol-Meyers Squibb. *A Phase III, multi-center, randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of BMS-188667 in subjects with active rheumatoid arthritis on background DMARDS who have failed anit-TNF therapy*. Clinical Study Report IM01029 2004. URL: http://ctr.bms.com/pdf//IM101029LT.pdf (accsessed October 2009).

132. Erratum: Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition (*N Engl J Med* 2005;**353**:1114–23). *N Engl J Med* 2005;**353**:2311.

133. Furst DE, Gaylis N, Bray V, Olech E, Yocum D, Ritter J, *et al.* Open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept: the opposite study. *Ann Rheum Dis* 2007;**66**:893–9.

134. Weinblatt M, Schiff M, Goldman A, Kremer J, Luggen M, Li T, *et al.* Selective costimulation modulation using abatacept in patients with active rheumatoid arthritis while receiving etanercept: a randomised clinical trial. *Ann Rheum Dis* 2007;**66**(2):228–34.

135. Weinblatt M, Combe B, Covucci A, Aranda R, Becker JC, Keystone E. Safety of the selective costimulation modulator abatacept in rheumatoid arthritis patients receiving background biologic and nonbiologic disease-modifying antirheumatic drugs: a one-year randomized, placebo-controlled study. *Arthritis Rheum* 2006;**54**:2807–16.

136. Finckh A, Ciurea A, Brulhart L, Kyburz D, Moller B, Dehler S, *et al.* B cell depletion may be more effective than switching to an alternative anti-tumor necrosis factor agent in rheumatoid arthritis patients with inadequate response to anti-tumor necrosis factor agents. *Arthritis Rheum* 2007;**56**:1417–23.

137. Finckh A, Ciurea A, Brulhart L, Moller B, Walker UA, Courvoisier D, *et al.* Which subgroup of rheumatoid arthritis patients benefits from switching to rituximab versus alternative anti-TNF agents after previous failure to anti-TNF agent? *Ann Rheum Dis* 2010;**69**:387–93.

138. Kievit W, Adang EM, Fransen J, Kuper HH, Van De Laar MAFJ, Jansen TL, *et al.* The effectiveness and medication costs of three anti-tumour necrosis factor alpha agents in the

treatment of rheumatoid arthritis from prospective clinical practice data. *Ann Rheum Dis* 2008;**67**:1229–34.

139. Roche. MabThera® (rituximab) for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor. *Submission to NICE*, 2009.

140. Triggiani M. Cost-effectiveness analysis of two biological treatments (Adalimumab vs. Etanercept) in moderate-severe rheumatoid arthritis. *Ital J Allergy Clin Immunol* 2006;**16**:47–50.

141. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;**313**:275–83.

142. Vera-Llonch M, Massarotti E, Wolfe F, Shadick N, Westhovens R, Sofrygin O, *et al.* Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to tumor necrosis factor-alpha antagonists. *J Rheumatol* 2008;**35**:1745–53.

143. Russell A, Beresniak A, Bessette L, Haraoui B, Rahman P, Thorne C, *et al.* Cost-effectiveness modeling of abatacept versus other biologic agents in DMARDS and anti-TNF inadequate responders for the management of moderate to severe rheumatoid arthritis. *Clin Rheumatol* 2009;**28**:403–12.

144. Kielhorn A, Porter D, Diamantopoulos A, Lewis G. Uk cost-utility analysis of rituximab in patients with rheumatoid arthritis that failed to respond adequately to a biologic disease-modifying antirheumatic drug. *Curr Med Res Opin* 2008;**24**:2639–50.

145. Lindgren P, Geborek P, Kobelt G. Modeling the cost-effectiveness of treatment of rheumatoid arthritis with rituximab using registry data from Southern Sweden. *Int J Technol Assess Health Care* 2009;**25**:181–9.

146. Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. *Ann Rheum Dis* 2005;**64**:995–1002.

147. Dixon WG, Hyrich KL, Watson KD, Lunt M, Galloway J, Ustianowski A, *et al.* Drug-specific risk of tuberculosis in patients with rheumatoid arthritis treated with anti-TNF therapy: results from the British Society for Rheumatology Biologics Register (BSRBR). *Ann Rheum Dis* 2010;**69**:522–8.

148. Geborek P, Crnkic M, Petersson IF, Saxne T. Etanercept, infliximab, and leflunomide in established rheumatoid arthritis: clinical experience using a structured follow up programme in southern Sweden. *Ann Rheum Dis* 2002;**61**:793–8.

149. Ducornau P, Kielhorn A, Wintfeld N. Comparison of linear and non-linear utility mapping between HAQ and EQ-5D using pooled data from the tolicizumab trials OPTION and LITHE. Annual Meeting of the British Society of Rheumatology (BSR), Glasgow, 2009.

150. Abbott. Adalimumab, etanercept, infliximab, abatacept and rituximab for the treatment of rheumatoid arthritis after failure of an anti-TNF agent. *Submission to NICE*, 2009.

151. Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. *Rheumatology* 2004;**43**:62–72.

152. National Coordinating Centre for Health Technology Assessment (NCCHTA). *Abatacept for the treatment of refractory rheumatoid arthritis, HTA ref 06/52/01, Evidence Review Group Report for NICE (Project) (Project record)*. Southampton: National Coordinating Centre for Health Technology Assessment (NCCHTA); 2007.

153. Boggs R, Sengupta N, Ashraf T. Estimating health utility from a physical function assessment in rheumatoid arthritis patients treated with adalimumab. *ISPOR* meeting presentation, Rotterdam, the Nederlands, 2002.

154. Davies A, Cifaldi MA, Segurado OG, Weisman MH. Cost-effectiveness of sequential therapy with tumor necrosis factor antagonists in early rheumatoid arthritis. *J Rheumatol* 2009;**36**:16–25.

155. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997;**36**:551–9.

156. Hawthorne G, Richardson J, Day N, Osborne R, McNeil H. *Construction and utility scaling of the assessment of quality of life (AQOL) instrument*. Melbourne, VIC: Centre for Health Program Evaluation; 2000.

157. Michaud K, Wolfe F. *EQ5D changes rheumatoid arthritis (RA) quality of life in United States: study of 11,289 RA patients*. Society for Medical Decision Making annual meeting, plenary. San Franciso, CA, 2005.

158. Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Stat Med* 2007;**26**:1237–54.

159. Wailoo A. *The sequential use of TNF-alpha inhibitors. Update to a report by the Decision Support Unit. Evaluations*. URL: www.nice.org.uk/guidance/index.jsp?action=folder&o=40507

160. Bristol-Myers Squibb. Abatacept (Orencia®) for the treatment of rheumatoid arthritis. *Submission* to NICE, 2007.

161. Liverpool Reviews and Implementation Group. Abatacept for the treatment of rheumatoid arthritis. *ERG report for NICE*, 2007.

162. Song F, Loke YK, Walsh T, Glenny A-M, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluation healthcare interventions: survey of published systematic reviews. *BMJ* 2009;**338**:b1147.

163. Hyrich KL, Lunt M, Davies R, Harrison M, Watson K, Symmons D. The British society for rheumatology biologics register: updated analysis on sequential use of anti-TNF therapies in patients with rheumatoid arthritis. *Submission to NICE*, 2009;1–30.

164. Wolfe F, Rasker JJ, Boers M, Wells GA, Michaud K. Minimal disease activity, remission, and the long-term outcomes of rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2007;**57**:935–42.

165. Ducournau P, Kielhorn A, Wintfeld N. *Comparison of linear and non-linear utility mapping between HAQ and EQ-5D using pooled data from the tocilizumab trials OPTION and LITHE*. Annual Meeting of the British Society of Rheumatology (BSR) Glasgow, 2009; poster 258.

166. Jacobsson LT, Turesson C, Nilsson J-A. Treatment with TNF blockers and mortality risk in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;**66**:670–5.

167. Moots RJ, Riel PV, Kekow J. Patterns of dose escalation and DMARD intensification in 739 patients with rheumatoid arthritis (RA) treated with anti-TNF agents (ATAS): results from the DART study. *Rheumatology* 2008; Abstract 418.

168. Krishnan E, Sokka T, Hakkinen A, Hubert H, Hannonen P. Normative values for the health assessment questionnaire disability index. *Arthritis Rheum* 2004;**50**:953–60.

169. Taylor M, Saxby R, Ganderton M, Conway P, Lebmeier P, Peckham J. *The relationship between Health Assessment Questionnaire score and resource use in the management of rheumatoid arthritis*. European League Against Rheumatism (EULAR), Paris, 2008.

170. Brennan A, Bansback N, Nixon R, Madan J, Harrison M, Watson K, *et al.* Modelling the cost effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. *Rheumatology* 2007;**46**:1345–54.

171. Schering-Plough Limited. Remicade® for the treatment of rheumatoid arthritis after the failure of a TNFá inhibitor in England and Wales. *Submission to NICE*, 2009.

172. Bansback N, Marra C, Tsuchiya A, Anis A, Guh D, Hammond T, *et al.* Using the health assessment questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis. *Arthritis Rheum* 2007;**57**:963–71.

173. Brennan A, Madan J, Bansback N, Nixon R, Symmoins D, Lunt M. *Modelling the cost-effectiveness of sequential use of TNF-alpha inhibitors in the management of rheumatoid arthritis: an update*. HEDS Discussion Paper Series 6/12; 2006.

174. Strangfeld A, Eveslage M, Kekow J, Gräßler A, Kaufmann J, Listing J, *et al.* Effectiveness of treatment with Rituximab depends on autoantibody status – results from 2 years of experience in the German Biologics Register RABBIT, ACR Conference, Philadelphia, PA, 2009.

175. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. *Health Technol Assess* 2004;**8**(11).

176. Ariza-Ariza R, Navarro-Sarabia F, Hernandez-Cruz B, Rodriguez-Arboleya L, Navarro-Compan V, Toyos J. Dose escalation of the anti-TNF- agents in patients with rheumatoid arthritis. A systematic review. *Rheumatology* 2007;**46**:529–32.

177. Wolfe F. The epidemiology of drug treatment failure in rheumatoid arthritis. *Baillieres Best Pract Res Clin Rheumatol* 1995;**9**:619–32.

178. National Institute for Health and Clinical Excellence. *Effect of a second course of anti-TNF therapy on HAQ following lack of response to the first course*. London: National Institute for Health and Clinical Excellence; 2008.

179. Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, *et al.* A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. *Health Technol Assess* 2006;**10**(42).

180. Fumagalli M, Incorvaia C, Nitti F. The assessment of quality of life as a measure of gold salts treatment efficacy in rheumatoid arthritis. *Minerva Med* 2002;**93**:199–202.

181. Poor G, Strand V. Efficacy and safety of leflunomide 10 mg versus 20 mg once daily in patients with active rheumatoid arthritis: multinational double-blind, randomized trial. *Rheumatology* 2004;**43**:744–9.

182. Keystone E, Fleischmann R, Emery P, Furst DE, Van VR, Bathon J, *et al.* Safety and efficacy of additional courses of rituximab in patients with active rheumatoid arthritis: an open-label extension analysis. *Arthritis Rheum* 2007;**56**:3896–908.

183. NIHR Coordinating Centre for Health Technology Assessment. *Rituximab for the treatment of rheumatoid arthritis. NICE Evidence Review Group Report (ERG) in support of NICE's Single Technology Appraisal process [1645]*. URL: www.nice.org.uk/nicemedia/live/11719/36074/36074.pdf (accessed November 2009).

184. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, *et al.* A trial of etanercept, a recombinant tumor necrosis factor receptor: Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;**340**:253–9.

185. Furst DE, Schiff MH, Fleischmann RM, Strand V, Birbara CA, Compagnone D, *et al.* Adalimumab, a fully human anti-tumor necrosis factor-alpha monoclonal antibody, and concomitant standard antirheumatic therapy for the treatment of rheumatoid arthritis: results of STAR (Safety Trial of Adalimumab in Rheumatoid Arthritis). *J Rheumatol* 2003;**30**:2563–71.

186. Westhovens R, Yocum D, Han J, Berman A, Strusberg I, Geusens P, *et al.* The safety of infliximab, combined with background treatments, among patients with rheumatoid arthritis and various comorbidities: a large, randomized, placebo-controlled trial. *Arthritis Rheum* 2006;**54**:1075–86.

187. Kruger K, Schattenkirchner M. Comparison of cyclosporin A and azathioprine in the treatment of rheumatoid arthritis – results of a double-blind multicentre study. *Clin Rheumatol* 1994;**13**:248–55.

188. van Rijthoven AW, Dijkmans BA, The HS. Comparison of cyclosporine and D-penicillamine for rheumatoid arthritis: a randomized, double blind, multicenter study. *J Rheumatol* 1991;**18**:815–20.

189. Liverpool Reviews and Implementation Group. *Erlotinib for the treatment of relapse non-small cell lung cancer*. Evidence Review Group (ERG) report, September 2006. URL: www.nice.org.uk/nicemedia/live/11714/35177/35177.pdf (accessed January 2011).

190. Wolfe F, Mitchell DM, Sibley JT. The mortality of rheumatoid arthritis. *Arthritis Rheum* 1994;**37**:481–94.

191. Curtis L. *Unit costs of community care. Personal Social Services Research Unit.* 2008. URL: www.pssru.ac.uk (accessed November 2009).

192. Barbieri M, Wong JB, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. *Pharmacoeconomics* 2005;**23**:607–18.

193. Pugner KM, Scott DI, Holmes JW, Hieke K. The costs of rheumatoid arthritis: an international long-term view. *Semin Arthritis Rheum* 2000;**29**:305–20.

194. Wolfe F, Zwillich SH. The long-term outcomes of rheumatoid arthritis: a 23-year prospective, longitudinal study of total joint replacement and its predictors in 1,600 patients with rheumatoid arthritis. *Arthritis Rheum* 1998;**41**:1072–82.

195. Hamilton P, Lemon M, Field R. Cost of total hip and knee arthroplasty in the UK. A comparison with the current reimbursement system in the NHS. *J Bone Joint Surg Br* 2009;**91-B**(112-a).

196. NICE Decision Support Unit. *Sequential use of TNF-a inhibitors for the treatment of rheumatoid arthritis*. URL: www.nice.org.uk/nicemedia/pdf/DSUReportWith AppendicesForConsultation.pdf (accessed 11 June 2010,).

197. Scott DL, Garrood T. Quality of life measures: use and abuse. *Baillieres Best Pract Res Clin Rheumatol* 2000;**14**:663–87.

198. Dixon W, Hyrich K, Watson K, Lunt M. Drug-specific risk of tuberculosis in patients with rheumatoid arthritis treated with anti-TNF therapy: results from the BSRBR. *American College of Rheumatology Annual Meeting* 2008; Presentation No: 1263.

199. Wolfe F, Mitchell DM, Sibley JT. The mortality of rheumatoid arthritis. *Arthritis Rheum* 1994;**37**:481–94.

200. Emery P, Breedveld FC, Lemmel EM, Kaltwasser JP, Dawes PT, Gomor PA. A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;**39**:655–65.

201. Hamilton J, McInnes IB, Thomson EA, Porter D, Hunter JA, Madhok R, *et al.* Comparative study of intramuscular gold and methotrexate in a rheumatoid arthritis population from a socially deprived area. *Ann Rheum Dis* 2001;**60**:566–72.

202. Yocum DE, Allard S, Cohen SB, Emery P, Flipo RM, Goobar J, *et al.* Microemulsion formulation of cyclosporin (Sandimmun Neoral) vs Sandimmun: comparative safety, tolerability and efficacy in severe active rheumatoid arthritis. On behalf of the OLR 302 Study Group. *Rheumatology* 2000;**39**:156–64.

203. Marra CA, Esdaile JM, Guh D, Fisher JH, Chalmers A, Anis AH. The effectiveness and toxicity of cyclosporin A in rheumatoid arthritis: longitudinal analysis of a population-based registry. *Arthritis Rheum* 2001;**45**:240–5.

204. Willkens RF, Sharp JT, Stablein D, Marks C, Wortmann R. Comparison of azathioprine, methotrexate, and the combination of the two in the treatment of rheumatoid arthritis. A forty-eight-week controlled clinical trial with radiologic outcome assessment. *Arthritis Rheum* 1995;**38**:1799–806.

205. Bristol-Myers Squibb Pharmaceuticals Ltd. Abatacept (Orencia) for the Treatment of Rheumatoid Arthritis. *Submission to NICE*, 2009.

206. Edwards CJ, Arden NK, Fisher DSJC, Reading I, van Staa TP. The changing use of disease-modifying anti-rheumatic drugs in individuals with rheumatoid arthritis from the United Kingdom General Practice Research Database. *Rheumatology* 2005;**44**:1398.

207. National Audit Office. *Survey of general practitioners about the diagnosis and management of rheumatoid arthritis*. URL: www.nao.org.uk/idoc.ashx?docId=de0440fa-a1a9-4d6c-9975-f04a74011030&version=-1 (accessed November 2009).

208. Royal College of Physicians. T*eams without walls. The value of medical innovation and leadership*. URL: www.rcplondon.ac.uk/professional-Issues/Documents/teams-without-walls.pdf (accessed November 2009).

209. Department of Health. *High quality care for all: NHS next stage review final report*. URL: www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/DH_085825 (accessed November 2009).

210. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcomes in arthritis. *Arthritis Rheum* 1980;**23**:137–45.

211. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:727–35.

212. Siegel JN, Zhen B-G. Use of the American College of Rheumatology N (ACR-N) index of improvement in rheumatoid arthritis: argument in favour. *Arthritis Rheum* 2005;**52**:1637–41.

213. Fransen J, Van Riel PL. The Disease Activity Score and the EULAR response criteria. *Clin Exp Rheumatol* 2005;**23**:S93–S99.

214. Scott DL. Radiographs in rheumatoid arthritis. *Int J of Adv Rheumatol* 2003;**1**:2–8.

215. Sharp JT, Young DY, Bluhm GB, Brook A, Brower AC, Corbett M. How many joints in the hands and wrists should be included in a score of radiologic abnormalities use to assess rheumatoid arthritis? *Arthritis Rheum* 1985;**28**:26–34.

216. Van Der Heijde D, van Leeuwen MA, Riel PLCM, Koster AM, van T Hof MA, van Rijswijk MH. Biannual radiographic assessments of hands and feet in a three-year prospective followup of patients with early rheumatoid arthritis. *Arthritis Rheum* 1992;**35**:26–34.

217. Scott DL, Housseien DA, Laasonen L. Proposed modification to Larsen's scoring methods of hand and wrist radiographs. *Rheumatology* 1995;**34**:56.

218. The United Kingdom Government Actuary's Department. GAD, 2005.

219. Nuijten MJ, Engelfriet P, Duijn K, Bruijn G, Wierz D, Koopmanschap M. A cost-cost study comparing etanercept with infliximab in rheumatoid arthritis. *Pharmacoeconomics* 2001;**19**:1051–64.

220. Kobelt G, Eberhardt K, Jönsson L, Jönsson B. Economic consequences of the progression of rheumatoid arthritis in Sweden. *Arthritis Rheum* 1999;**42**:347–56.

221. Kobelt G, Lindgren P, Singh A, Klareskog L. Cost effectiveness of etanercept (Enbrel) in combination with methotrexate in the treatment of active rheumatoid arthritis based on the TEMPO trial. *Ann Rheum Dis* 2005;**64**:1174–9.

222. National Institute for Health and Clinical Excellence. Guidance on the use of etanercept and infliximab for the treatment of rheumatoid arthritis. Technology Appraisal no. 36, 2002.

223. Hyrich KL, Silman AJ, Lunt M. Influence of response and adverse events rates to a first anti-TNF alpha agent on the outcome from switching to a second agent: results from British Society for Rheumatology Biologics Register. BSR Biologics Register Newsletter, 2006.

224. Kielhorn A, Tony H, Jost F, Adultman R. Rituximab in rheumatoid arthritis: translating ACR responses into benefit for the patients. *Ann Rheum Dis* 2006;**65**:324.

225. Scott DL, Garrood T. Quality of life measures: use and abuse. *Bailliere's Best Prac Res Clin Rheumatol* 2000;**14**:663-87.

226. Wyeth Pharmaceuticals. Etanercept (ENBREL). Submission to NICE, 2009.

227. Hetland ML, Chistensen JL, Tarp U, Dreyer L, Hansen A, Hansen IT, *et al.* Direct comparison of 4 years' drug survival of adalimumab, etanercept and infliximab in rheumatoid arthritis patients. An observational study from the DANBIO Registry. Presentation 997, ACR/ARHP Scientific meeting, Philadelphia, PA, 2009.

228. Curtis JR, John A, Baser O. Determinants of tumor necrosis factor alpha inhibitor switching in patients with rheumatoid arthritis. ACR/ARHP Scientific meeting, Philadelphia, PA 2009; presentation 1004.

229. Zink A, Listing J, Kary S, Ramlau P, Stoyanova-Scholz M, Babinsky K, *et al.* Treatment continuation in patients receiving biological agents or conventional DMARD therapy. *Ann Rheum Dis* 2005;**64**:1274–9.

230. Wolfe F, Michaud K. Duration of use of anti-TNF therapy in rheumatoid arthritis. ACR/ARHP Scientific meeting, Boston, MA 2007; presentation 969.

231. Finckh A, Simard JF, Gabay C, Guerne P-A. Evidence for differential acquired drug resistance to anti-tumour necrosis factor agents in rheumatoid arthritis. *Ann Rheum Dis* 2006;**65**:746–52.

232. Duclos M, Gossec L, Ruyssen-Witrand A, Salliot C, Luc M, Guignard S, *et al.* Retention rates of tumor necrosis factor blockers in daily practice in 770 rheumatic patients. *J Rheumatol* 2006;**33**:2433–9.

233. Van Vollenhoven RF, Carli CC, Bratt J, Klareskog L. Secondary loss of efficacy with TNF-alpha antagonists. Data from the STURE registry. *Rheumatology* 2005;**52**:275.

234. Kristensen LE, Saxne T, Nilsson JA, Geborek P. Impact of concomitant DMARD therapy on adherence to treatment with etanercept and infliximab in rheumatoid arthritis. Results from a six-year observational study in southern Sweden. *Arthritis Res Ther* 2006;**8**:R174.

Appendix 1

Details of key outcomes used in rheumatoid arthritis trials

The Health Assessment Questionnaire

The HAQ now comprises a family of questionnaires designed to assess the functional capacity of patients with musculoskeletal complaints and specifically RA. The most widely used HAQ is derived from the Stanford Health Assessment Questionnaire[210] and consists of two or three questions in eight categories:

- dressing and grooming: dress yourself, including doing shoelaces, and shampooing your hair
- rising: from an armless chair and in and out of bed
- eating: being able to cut meat, lift a full cup or glass to mouth, and open a new carton of milk
- walking: outdoors on flat ground and climb five steps
- hygiene: wash and dry entire body, take a bath, get on and off the toilet
- reaching: reach and get down a 5-lb object, bend down and pick up clothing
- grip: open car doors, open previously unopened jars, turn taps on and off
- activities: run errands and shop, get in and out of car, do chores.

The score from the most limited activity in each category is obtained. Each category is scored 0 (without any difficulty), 1 (with some difficulty), 2 (with much difficulty) or 3 (unable to do). Use of aids or devices to help with function is taken into account so that need for such assistance automatically scores 2 (unless 3 has been ticked). The maximum score in each of the eight categories is added to give a maximum possible score of 24. This total score may be divided by 8 to give an average value in the range 0–3.

The HAQ has several modifications:[39]

- Modified HAQ (MHAQ) is a shortened version of HAQ which uses only one question in each of the eight categories and does not consider the use of aids and devices to assist function. It is simpler to score and has the same range as HAQ (0–3).
- RA-HAQ is another shortened version of HAQ designed to overcome some of the metric limitations of MHAQ.
- DHAQ this uses the original eight categories of HAQ, but is based on the most difficult items in each of the categories. Neither the RA-HAQ nor DHAQ have been widely used, unlike MHAQ.

American College for Rheumatology response criteria[209]

In order to achieve an ACR20 response a 20% improvement in the score for tender joints and a 20% improvement in swollen joints is necessary and 20% improvement in at least three of the following:

- global disease activity assessed by observer
- global disease activity assessed by patient
- patient assessment of pain
- physical disability score (e.g. HAQ)
- acute phase response (e.g. ESR or CRP).

Responses may also be defined as ACR50 (50%) or ACR70 (70%) depending on degree of benefit.

The ACR-N is an extension of the ACR response criteria, and is defined as the lowest of the following three values:

- percentage change in the number of swollen joints
- percentage change in the number of tender joints
- the median of the percentage change in the other five measures listed above.

It is thus a continuous variable. For example, an ACR-N score of 38 means an improvement of at least 38% in tender joint counts (TJCs) and swollen joint counts (SJCs) and an improvement of at least 38% in three of the five other parameters.[212]

DAS

Original DAS

DAS = 0.54($\sqrt{}$RAI[a]) + 0.065(total number of swollen joints out of 44) + 0.33(In ESR[b]) + 0.0072 (patient general health score where 0 = best, 100 = worst).

(a) RAI refers to a graded score of joint tenderness for 53 joints known as the Ritchie Articular Index and (b) the ESR.

DAS based on 28 joint evaluations

DAS28 – 4 = 0.56($\sqrt{}$TJC28) + 0.28($\sqrt{}$SJC28) + 0.7ln(ESR) + 0.014(patient general health score where 0 = best, 100 = worst).

Where scores for general health are not available, or not measured, the following formula is used:

$$DAS28 - 3 = [0.56(\sqrt{}TJC28) + 0.28(\sqrt{}SJC28) + 0.7\ln(ESR)]1.08 + 0.16$$

EULAR response criteria

The EULAR response criteria[213] are based on the DAS score. They incorporate both change from baseline and DAS or DAS28 at end point and, based on both, classify patients as good or moderate responders or non-responders (*Table 107*).

Radiographic assessment methods[212]

Sharp Score

The simplified Sharp system,[215] which evaluates hand and wrist images, assesses 17 areas for erosions and 18 areas for joint space narrowing. Each joint is scored on a 6-point scale as follows: 0 = no erosion; 1 = discrete erosion; 2 = two separate quadrants with erosions or 20%–40% joint involvement; 3 = three separate quadrants with erosions or 41%–60% joint involvement; 4 = all four quadrants with joint erosion or 61%–80% joint involvement; and 5 = extensive destruction with greater than 80% joint involvement. The range of erosion scores for a patient with two hands and wrists is 0–170. For joint space narrowing each joint is scored using a 5-point scale

TABLE 107 The EULAR response criteria using DAS and DAS28

		Improvement in DAS or DAS28 from baseline		
DAS at end point	DAS28 at end point	≥1.2	>0.6 and ≤1.2	≤0.6
≤2.4	≤3.2	Good		
>2.4 and ≤3.7	>3.2 and ≤5.1		Moderate	
>3.7	>5.1			None

as follows: 0 = no narrowing; 1 = up to 25% narrowing; 2 = 26%–65% narrowing; 3 = 66%–99% narrowing; 4 = complete narrowing. The range for joint space narrowing is therefore 0–144. This gives a total joint score in the range 0–314.

Van der Heijde-modified Sharp score

In this case 16 joints are assessed in each hand and wrist and six joints in each foot. Erosions are scored 0–5 and depending on the affected surface area and 0–10 in the fee, yielding possible erosion scores of 0–160 for hands/wrists and 0–120 for feet (total 0–280). Joint space narrowing is assessed in 15 joints for each hand/wrist and six joints in each foot on a scale of 0–4. The range of possible joint space narrowing scores is in the range 0–168. This yields a possible total score in the range 0– 448.[216]

The Larsen score

In this method standard films are used to classify each joint into one of six possible categories (0 = normal, 5 = severely damaged). Any joint may be scored but the focus is on hands and feet. In the hands each proximal interphalangeal joint and each metacarpophalangeal joint scores 0–5; each wrist joint scores 0–25 (the basic score is multiplied by 5): this gives a maximum score of 150 for two hands and wrists. In the feet each metatarsophalangeal joint is scored 0–5, giving a total score of 50 for two feet. This yields a possible total score in the range 0–200.

Scott-modified Larsen

Scott *et al.*[217] suggested minor modifications to the scale in order to improve correlation between scorers. It was proposed that grade 1 included erosions and cysts of less than 1 mm diameter and grade included one or more erosions of greater than 1 mm diameter.

Appendix 2

Literature search strategies

Source – Cochrane Library (CENTRAL, DARE and NHS EED) 2009 Issue 3

#1 rheumatoid next arthritis
#2 MeSH descriptor Arthritis, Rheumatoid explode all trees
#3 (#1 OR #2)
#4 adalimumab or humira
#5 etanercept or enbrel
#6 infliximab or remicade
#7 rituximab or mabthera
#8 abatacept or orencia
#9 (#4 OR #5 OR #6 OR #7 OR #8)
#10 (#3 AND #9)

Source – MEDLINE (Ovid) 1950 – July Week 1 2009

1. rheumatoid arthritis.tw. (58,668)
2. arthritis rheumatoid/ (68,937)
3. or/1-2 (83,478)
4. (adalimumab or humira).mp. (1,199)
5. (etanercept or enbrel).mp. (2,138)
6. (rituximab or mabthera).mp. (5,052)
7. (abatacept or orencia).mp. (1,779)
8. (infliximab or remicade).mp. (4,830)
9. or/4-8 (13,083)
10. 3 and 9 (2,759)

Source – MEDLINE(Ovid) In-Process & Other Non-Indexed Citations 13 July, 2009

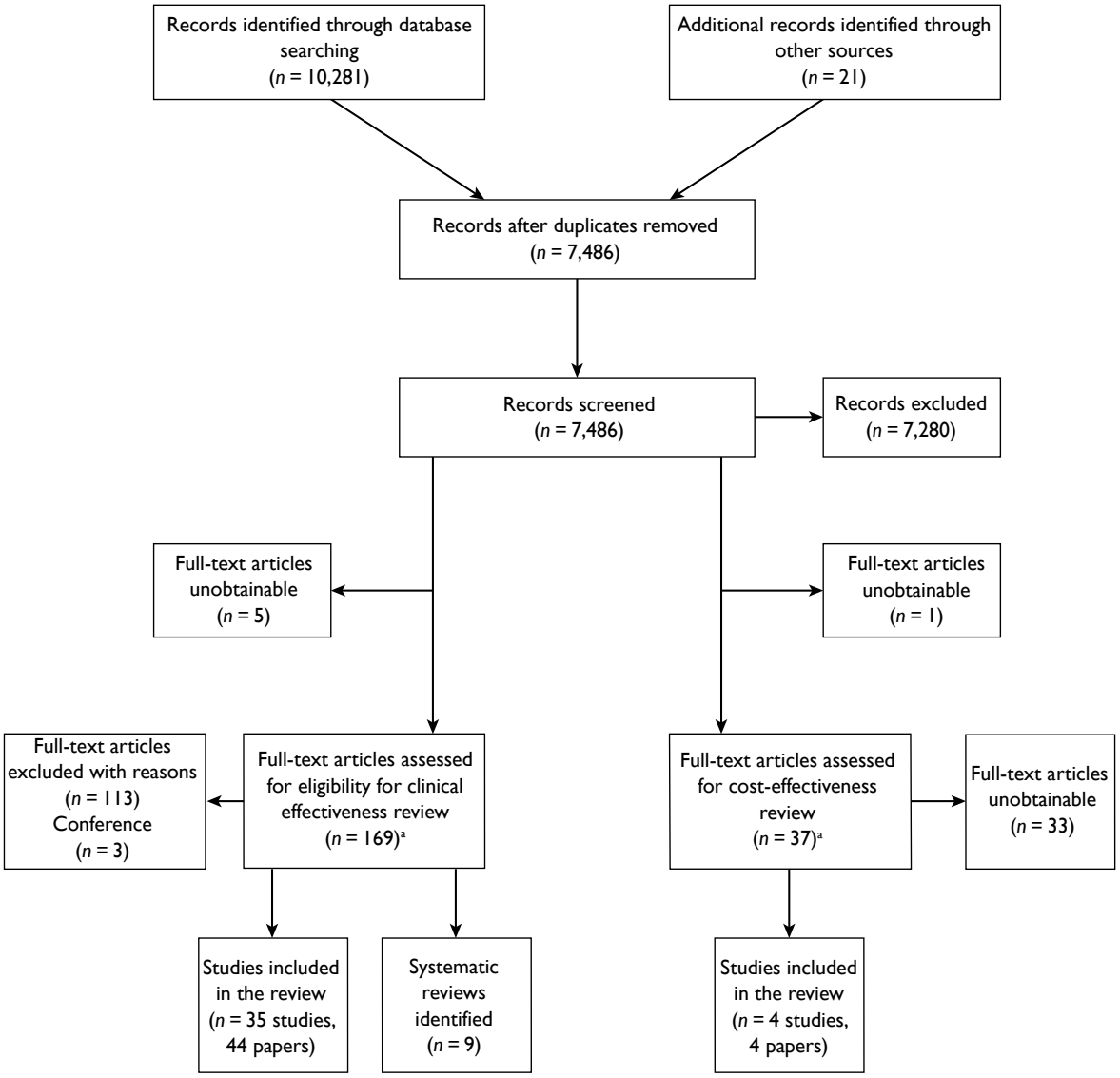
11. (adalimumab or humira).mp. (129)
12. (etanercept or enbrel).mp. (203)
13. (rituximab or mabthera).mp. (455)
14. (abatacept or orencia).mp. (39)
15. (infliximab or remicade).mp. (346)
16. or/1-5 (990)
17. rheumatoid arthritis.tw. (1,987)
18. 6 and 7 (220)

Source – EMBASE (Ovid) 1980 to 2009 Week 28

19. (adalimumab or humira).ti,ab,sh. (4,120)
20. (etanercept or enbrel).ti,ab,sh. (8,362)
21. (rituximab or mabthera).ti,ab,sh. (12,634)
22. (abatacept or orencia).ti,ab,sh. (1,014)
23. (infliximab or remicade).ti,ab,sh. (12,117)
24. or/1-5 (26,879)
25. rheumatoid arthritis/ (59,837)
26. rheumatoid arthritis.tw. (47,871)
27. 7 or 8 (68,003)
28. 6 and 9 (6,262)

Appendix 3

Flow diagram



[a]One paper was ordered for both clinical effectiveness and cost-effectiveness.

Appendix 4

Clinical effectiveness: table of excluded studies with rationale

Article	Reason for exclusion
Prior lack of efficacy with etanercept does not predict lack of efficacy with infliximab. *Formulary* 2005;**40**:93.	Design
Abatacept: rheumatoid arthritis: after failure of TNF alpha antagonists and rituximab. *Prescrire Int* 2008;**17**:232.	Design
[Fusion protein abatacept. Remission in every 5th TNF-alpha refractory patient]. [German]. *MMW Fortschritte der Medizin* 2008;**150**:56–7.	Design
The COMET study: high remission rate through the use of etanercept in early rheumatoid arthritis. [German]. *Arzneimitteltherapie* 2008;**26**:434–5.	Population
Alexander W, Han C, Giles J. American College of Rheumatology Scientific Meeting. ASPIRE: Infliximab (Remicade) plus methotrexate for rheumatoid arthritis. *P T* 2009;**34**:37.	Population
Allison C. Abatacept as add-on therapy for rheumatoid arthritis. *Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA)* 2005;**4**.	Design
Allison C. Abatacept as add-on therapy for rheumatoid arthritis. *Issues Emerg Health Technol* 2005;**73**:1–4.	Design
Alonso-Ruiz A, Pijoan JI, Ansuategui E, Urkaregi A, Calabozo M, Quintana A. Tumour necrosis factor alpha drugs in rheumatoid arthritis: systematic review and metaanalysis of efficacy and safety. *BMC Musculoskelet Disord* 2008;**9**:52.	Population
Alten R. Costimulation with abatacept – a new and successful therapeutic principle in rheumatoid athritis. Part 2: efficacy and safety of abatacept. *Aktuelle Rheumatologie* 2007;**32**:271–7.	Design
Alten R, Musch A. Abatacept in patients with rheumatoid arthritis. *Arzneimitteltherapie* 2008;**26**:9–16.	Design
Arenere MM, Navarro AH, Cilveti SU, Allende BM, Rabanaque HM, Arrieta NR, *et al.* Etanercept use in rheumatoid arthritis patients treated previously with infliximab. *Atencion Farmaceutica* 2005;**7**:465–9.	Participant number
Assous N, Gossec L, Dougados M, Kahan A, Allanore Y. Efficacy of rituximab in patients with rheumatoid arthritis refractory or with contra-indication to anti-tumour necrosis factor-alpha drugs in daily practice: an open label observational study. *Clin Exp Rheumatol* 2007;**25**:504.	Participant number
Baumgartner SW, Fleischmann RM, Moreland LW, Schiff MH, Markenson J, Whitmore JB. Etanercept (Enbrel) in patients with rheumatoid arthritis with recent onset versus established disease: improvement in disability. *J Rheumatol* 2004;**31**:1532–7.	Population
Bazzani C, Filippini M, Caporali R, Bobbio-Pallavicini F, Favalli EG, Marchesoni A, *et al.* Anti-TNFalpha therapy in a cohort of rheumatoid arthritis patients: clinical outcomes. *Autoimmun Rev* 2009;**8**:260–5.	Population
Bernal RL, Guerrero A, Monzon MA, Beltran GM, Hernandez CB, Colmenero MA. Effectiveness and safety of adalimumab and etanercept for rheumatoid arthritis in a third-level hospital. *Farm Hosp* 2006;**30**:223–9.	Population
Blank N, Max R, Schiller M, Briem S, Lorenz H-M. Safety of combination therapy with rituximab and etanercept for patients with rheumatoid arthritis. *Rheumatology* 2009;**48**:440–1.	Participant number
Blumenauer Barbara BTB, Judd M, Wells GA, Burls A, Cranney A, Hochberg MC, *et al. Infliximab for the treatment of rheumatoid arthritis. Reviews.* Cochrane Database of Systematic Reviews 2002 Issue 3. Chichester (UK): John Wiley and Sons, Ltd; 2002.	Population
Blumenauer Barbara BTB, Cranney A, Burls A, Coyle D, Hochberg MC, Tugwell P, *et al. Etanercept for the treatment of rheumatoid arthritis.* Cochrane Database of Systematic Reviews 2003 Issue 3. Chichester (UK): John Wiley and Sons, Ltd; 2003.	Population
Braun-Moscovici Y, Markovits D, Rozin A, Toledano K, Nahir AM, Balbir-Gurman A. Anti-tumour necrosis factor therapy: 6 year experience of a single centre in northern Israel and possible impact of health policy on results. *Isr Med Assoc J* 2008;**10**:277–81.	Population
Brocq O, Plubel Y, Breuil V, Grisot C, Flory P, Mousnier A, *et al.* Etanercept – infliximab switch in rheumatoid arthritis 14 out of 131 patients treated with anti TNFalpha. *Presse Med* 2002;**31**:1836–9.	Participant number
Brocq O, Plubel Y, Breuil V, Grisot C, Flory P, Mousnier A, *et al.* Switch etanercept – infliximab dans la polyarthrite rhumatoide. *Presse Med* 2002;**31**:1836–9.	Participant number
Brocq O, Albert C, Roux C, Gerard D, Breuil V, Ziegler LE. Adalimumab in rheumatoid arthritis after failed infliximab and/or etanercept therapy: experience with 18 patients. *Joint Bone Spine* 2004;**71**:601–3.	Participant number

Article	Reason for exclusion
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* 2004;**43**:243–4.	Population
Buch MH, Bingham SJ, Seto Y, McGonagle D, Bejarano V, White J, *et al*. Lack of response to Anakinra in rheumatoid arthritis following failure of tumour necrosis factor alpha blockade. *Arthritis Rheum* 2004;**50**:725–8.	Intervention
Buch MH, Boyle DL, Rosengren S, Saleem B, Reece RJ, Rhodes LA, *et al*. Mode of action of abatacept in rheumatoid arthritis patients having failed tumour necrosis factor blockade: a histological, gene expression and dynamic magnetic resonance imaging pilot study. *Ann Rheum Dis* 2009;**68**:1220–7.	Participant number
Burmester GR, Mariette X, Montecucco C, Monteagudo-Saez I, Malaise M, Tzioufas AG, *et al*. Adalimumab alone and in combination with disease-modifying antirheumatic drugs for the treatment of rheumatoid arthritis in clinical practice: the Research in Active Rheumatoid Arthritis (ReAct) trial. *Ann Rheum Dis* 2007;**66**:732–9.	Population
Burr ML, Malaviya AP, Gaston JH, Carmichael AJ, Ostor AJK. Rituximab in rheumatoid arthritis following anti-TNF-associated tuberculosis. *Rheumatology* 2008;**47**:738–9.	Design
Carmona L. Changes in anti-TNF: is this always justified? *Reumatologia Clinica* 2008;**4**:87–9.	Design
Combe B. Switching between anti-TNFalpha agents: what is the evidence? *Joint Bone Spine* 2004;**71**:169–71.	Design
Coyle D, Judd M, Blumenauer B, Cranney A, Maetzel A, Tugwell P, *et al*. Infliximab and etanercept in patients with rheumatoid arthritis: a systematic review and economic evaluation (DARE structured abstract). *Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA)* 2006;**45**.	Population
Davies A, Cifaldi MA, Segurado OG, Weisman MH. Cost-effectiveness of sequential therapy with tumour necrosis factor antagonists in early rheumatoid arthritis. *J Rheumatol* 2009;**36**:16–25.	Design
Di PE, Perin A, Morassi MP, Del FM, Ferraccioli GF, De VS. Switching to etanercept in patients with rheumatoid arthritis with no response to infliximab. *Clin Exp Rheumatol* 2007;**25**:85–7.	Participant number
Donahue KE, Gartlehner G, Jonas DE, Lux LJ, Thieda P, Jonas BL, *et al*. Systematic review: comparative effectiveness and harms of disease-modifying medications for rheumatoid arthritis. *Ann Intern Med* 2008;**148**:124–34.	Population
Emery P. Abatacept has beneficial effects in rheumatoid arthritis patients with an inadequate response to anti-TNFalpha therapy. *Clin Exp Rheumatol* 2005;**23**:767–8.	Design
Emery P, Fleischmann R, Filipowicz-Sosnowska A, Schechtman J, Szczepanski L, Kavanaugh A, *et al*. The efficacy and safety of rituximab in patients with active rheumatoid arthritis despite methotrexate treatment: results of a phase IIb randomised, double-blind, placebo-controlled, dose-ranging trial. *Arthritis Rheum* 2006;**54**:1390–400.	Population
Emery P, Keystone E, Tony HP, Cantagrel A, Van VR, Sanchez A, *et al*. IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. *Ann Rheum Dis* 2008;**67**:1516–23.	Intervention
Erickson AR, Mikuls TR. Switching anti-TNF-alpha agents: what is the evidence? *Curr Rheumatol Rep* 2007;**9**:416–20.	Design
Favalli EG, Arreghini M, Arnoldi C, Panni B, Marchesoni A, Tosi S, *et al*. Anti-tumour necrosis factor alpha switching in rheumatoid arthritis and juvenile chronic arthritis. *Arthritis Rheum* 2004;**51**:301–2.	Participant number
Fernandez Lison LC, Vazquez DB, Luis FJ, Moreno AP, Fruns GI, Liso RJ. Quality of life of patients with rheumatoid arthritis undergoing out-patient treatment with TNF inhibitors. *Farm Hosp* 2008;**32**:178–81.	Population
Filippini M, Bazzani C, Zingarelli S, Ziglioli T, Nuzzo M, Vianelli M, *et al*. Anti-TNF alpha agents in elderly patients with rheumatoid arthritis: a study of a group of 105 over sixty five years old patients. *Reumatismo* 2008;**60**:41–9.	Population
Gartlehner G, Hansen RA, Jonas BL, Thieda P, Lohr KN. The comparative efficacy and safety of biologics for the treatment of rheumatoid arthritis: a systematic review and metaanalysis. *J Rheumatol* 2006;**33**:2398–408.	Population
Genta MS, Kardes H, Gabay C. Clinical evaluation of a cohort of patients with rheumatoid arthritis treated with anti-TNF-alpha in the community. *Joint Bone Spine* 2006;**73**:51–6.	Population
Gomez-Puerta JA, Sanmarti R, Rodriguez-Cros JR, Canete JD. Etanercept is effective in patients with rheumatoid arthritis with no response to infliximab therapy. *Ann Rheum Dis* 2004;**63**:896.	Participant number
Gomez CT. Rituximab and abatacept in rheumatoid arthritis. *Reumatologia Clinica* 2009;**5**:77–81.	Design
Gonzalez-Juanatey C, Llorca J, Sanchez AA, Garcia-Porrua C, Martin J, Gonzalez-Gay MA. Short-term adalimumab therapy improves endothelial function in patients with rheumatoid arthritis refractory to infliximab. *Clin Exp Rheumatol* 2006;**24**:309–12.	Participant number
Gonzalez-Juanatey C, Llorca J, Vazquez-Rodriguez TR, az-Varela N, Garcia-Quiroga H, Gonzalez-Gay MA. Short-term improvement of endothelial function in rituximab-treated rheumatoid arthritis patients refractory to tumour necrosis factor alpha blocker therapy. *Arthritis Care Res (Hoboken)* 2008;**59**:1821–4.	Participant number
Haraoui B. Is there a rationale for switching from one anti-tumour necrosis factor agent to another? *J Rheumatol* 2004;**31**:1021–2.	Design
Hay EM, Thomas E, Paterson SM, Dziedzic K, Croft PR. Do etanercept-naqive patients with rheumatoid arthritis respond better to infliximab than patients for whom etanercept has failed? *Ann Rheum Dis* 2004;**63**:607–12.	Design

Article	Reason for exclusion
Health Q, I, Scotland. *Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis*. Glasgow: NHS Quality Improvement Scotland (NHS QIS); 2007.	Population
Health Q, I, Scotland. *Rituximab for the treatment of rheumatoid arthritis*. Glasgow: NHS Quality Improvement Scotland (NHS QIS); 2007.	Design
Heiberg MS, Rodevand E, Mikkelsen K, Kaufmann C, Didriksen A, Mowinckel P, *et al*. Adalimumab and methotrexate is more effective than adalimumab alone in patients with established rheumatoid arthritis: results from a 6-month longitudinal, observational, multicentre study. *Ann Rheum Dis* 2006;**65**:1379–83.	Population
Higashida J, Wun T, Schmidt S, Naguwa SM, Tuscano JM. Safety and efficacy of rituximab in patients with rheumatoid arthritis refractory to disease modifying antirheumatic drugs and anti-tumour necrosis factor-alpha treatment. *J Rheumatol* 2005;**32**:2109–15.	Population
Hoff M, Kvien TK, Kalvesten J, Elden A, Haugeberg G. Adalimumab therapy reduces hand bone loss in early rheumatoid arthritis: explorative analyses from the PREMIER study. *Ann Rheum Dis* 2009; **68**:1171–6.	Population
Iking-Konert C. Therapy-refractive rheumatoid arthritis: effectiveness and reliability of abatecept and infliximab. *Aktuelle Rheumatologie* 2008;**33**:239–40.	Population
Jamal S, Patra K, Keystone EC. Adalimumab response in patients with early versus established rheumatoid arthritis: DE019 randomised controlled trial subanalysis. *Clin Rheumatol* 2009;**28**:413–9.	Population
Kavanaugh A, Rosengren S, Lee SJ, Hammaker D, Firestein GS, Kalunian K, *et al*. Assessment of rituximab's immunomodulatory synovial effects (ARISE trial). 1: clinical and synovial biomarker results. *Ann Rheum Dis* 2008;**67**:402–8.	Participant number
Kielhorn A, Porter D, Diamantopoulos A, Lewis G. Uk cost-utility analysis of rituximab in patients with rheumatoid arthritis that failed to respond adequately to a biologic disease-modifying antirheumatic drug. *Curr Med Res Opin* 2008;**24**:2639–50.	Design
Kievit W, Adang EM, Fransen J, Kuper HH, Van De Laar MAFJ, Jansen TL, *et al*. The effectiveness and medication costs of three anti-tumour necrosis factor alpha agents in the treatment of rheumatoid arthritis from prospective clinical practice data. *Ann Rheum Dis* 2008;**67**:1229–34.	Population
Koike T, Harigai M, Inokuma S, Inoue K, Ishiguro N, Ryu J, *et al*. Postmarketing surveillance of the safety and effectiveness of etanercept in Japan. *J Rheumatol* 2009;**36**:898–906.	Population
Kristensen LE, Saxne T, Geborek P. The LUNDEX, a new index of drug efficacy in clinical practice: results of a five-year observational study of treatment with infliximab and etanercept among rheumatoid arthritis patients in Southern Sweden. *Arthritis Rheum* 2006;**54**:600–6.	Population
Laas K, Peltomaa R, Puolakka K, Kautiainen H, Leirisalo-Repo M. Early improvement of health-related quality of life during treatment with etanercept and adalimumab in patients with rheumatoid arthritis in routine practice. *Clin Exp Rheumatol* 2009;**27**:315–20.	Population
Li S, Kaur PP, Chan V, Berney S. Use of tumour necrosis factor-alpha (TNF-alpha) antagonists infliximab, etanercept, and adalimumab in patients with concurrent rheumatoid arthritis and hepatitis B or hepatitis C: a retrospective record review of 11 patients. *Clin Rheumatol* 2009;**28**:787–91.	Population
Lopez-Olivo MA, Amezaga M, McGahan L, Suarez-Almazor ME. Rituximab for rheumatoid arthritis. *Cochrane Database Syst Rev* 2008.	Design
Mease PJ, Revicki DA, Szechinski J, Greenwald M, Kivitz A, Barile-Fabris L, *et al*. Improved health-related quality of life for patients with active rheumatoid arthritis receiving rituximab – results of the dose-ranging assessment: international clinical evaluation of rituximab in rheumatoid arthritis (DANCER) trial. *J Rheumatol* 2008;**35**:20–30.	Population
Miyasaka N. Clinical investigation in highly disease-affected rheumatoid arthritis patients in Japan with adalimumab applying standard and general evaluation: the CHANGE study. *Mod Rheumatol* 2008;**18**:252–62.	Population
Moreland L. Efficacy of costimulation blockade with abatacept in rheumatoid arthritis patients refractory to tumour necrosis factor-alpha inhibition. *Curr Rheumatol Rep* 2006;**8**:367.	Design
Navarra SV, Raso A-A, Lichauco JJ, Tan PP. Clinical experience with infliximab among Filipino patients with rheumatic diseases. *APLAR J Rheumatol* 2006;**9**:150–6.	Population
Navarro-Sarabia F, riza-Ariza R, Hernandez-Cruz B, Villanueva I. *Adalimumab for treating rheumatoid arthritis*. Cochrane Database of Systematic Reviews 2005 Issue 3. Chichester (UK): John Wiley and Sons, Ltd; 2005.	Population
NHS Quality Improvement Scotland (NHS QIS). *Abatacept for the treatment of rheumatoid arthritis*. Glasgow: NHS Quality Improvement Scotland (NHS QIS); 2008.	Design
Nixon R, Bansback N, Brennan A. The efficacy of inhibiting tumour necrosis factor alpha and interleukin 1 in patients with rheumatoid arthritis: a meta-analysis and adjusted indirect comparisons. *Rheumatology* 2007;**46**:1140–7.	Population
Olsen N. Anti-TNF switching: effect on outcomes in patients with RA: commentary. *Nat Clin Pract Rheumatol* 2007;**3**:430–1.	Design
Ostergaard M, Unkerskov J, Linde L, Krogh NS, Ravn T, Ringsdal VS, *et al*. Low remission rates but long drug survival in rheumatoid arthritis patients treated with infliximab or etanercept: results from the nationwide Danish DANBIO database. *Scand J Rheumatol* 2007;**36**:151–4.	Population

Article	Reason for exclusion
Ostor AJK. Abatacept: a T-cell costimulation modulator for the treatment of rheumatoid arthritis. *Clin Rheumatol* 2008;**27**:1343–53.	Design
Owczarczyk KM, Hellmann M, Fliedner G, Rohrs T, Maizus K, Passon D, *et al*. Clinical outcome and B cell depletion in patients with rheumatoid arthritis receiving rituximab monotherapy in comparison with patients receiving concomitant methotrexate. *Ann Rheum Dis* 2008;**67**:1648–50.	Population
Palylyk-Colwell E, McGahan L. Rituximab for rheumatoid arthritis. *Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH)* 2006;**4**.	Design
Parker CT, Rennie T, Yocum DE, Furst DE, Kaine JL, Baldassare A, *et al*. Failure to report previously used drugs and dosages in pharmaceutical company-sponsored rheumatoid arthritis trials: comment on the article by Yocum *et al*. *Arthritis Rheum* 2004;**50**:3051–2.	Population
Pavelka K, Gatterova J, Vencovsky J, Sedova L, Chroust K. Radiographic progression of rheumatoid arthritis in real clinical practice results in national registry attra. *Rheumatologia* 2009;**23**:7–11.	Population
Pedersen SJ, Hetland ML, Ostergaard M, Navarro-Sarabia F, riza-Ariza R, Hernandez-Cruz B, *et al*. Adalimumab for treating rheumatoid arthritis. *Ugeskr Laeger* 2006;**168**:2899–902.	Population
Pisetsky DS. A landmark study on treatment strategies for rheumatoid arthritis. *Arthritis Rheum* 2008;**58**:S123–S125.	Population
Reynolds J, Shojania K, Marra CA. Abatacept: a novel treatment for moderate-to-severe rheumatoid arthritis. *Pharmacotherapy* 2007;**27**:1693–701.	Design
Rubbert-Roth A, Finckh A. Treatment options in patients with rheumatoid arthritis failing initial TNF inhibitor therapy: a critical review. *Arthritis Res Ther* 2009;**11**.	Design
Russell A, Beresniak A, Bessette L, Haraoui B, Rahman P, Thorne C, *et al*. Cost-effectiveness modelling of abatacept versus other biologic agents in DMARDS and anti-TNF inadequate responders for the management of moderate to severe rheumatoid arthritis. *Clin Rheumatol* 2009;**28**:403–12.	Design
Salliot C, Gossec L, Ruyssen-Witrand A, Luc M, Duclos M, Guignard S, *et al*. Infections during tumour necrosis factor-alpha blocker therapy for rheumatic diseases in daily practice: a systematic retrospective study of 709 patients. *Rheumatology* 2007;**46**:327–34.	Population
Sanmarti R, Gomez-Puerta JA, Rodriguez-Cros JR, Albaladejo C, Munoz-Gomez J, Canete JD. Etanercept in rheumatoid arthritis patients with a poor therapeutic response to infliximab. *Medicina Clinica* 2004;**122**:321–4.	Participant number
Sheitanov I. Our experience with Remicade (infliximab) in patients with early and refractory rheumatoid arthritis. *Rheumatology* 2005;**13**:66–73.	Population
Silman AJ. Available therapeutic options following failure of a first anti-TNF agent. *Nat Clin Pract Rheumatol* 2009;**5**:115.	Design
Singh A, Ghazvini P, Honeywell M, Treadwell P. Rituximab for the treatment of refractory rheumatoid arthritis: new information from clinical trials. *P T* 2006;**31**:321 + 343.	Design
Smolen JS, Weinblatt ME. When patients with rheumatoid arthritis fail tumour necrosis factor inhibitors: what is the next step? *Ann Rheum Dis* 2008;**67**:1497–8.	Design
Strand V, Balbir-Gurman A, Pavelka K, Emery P, Li N, Yin M, *et al*. Sustained benefit in rheumatoid arthritis following one course of rituximab: improvements in physical function over 2 years. *Rheumatology* 2006;**45**:1505–13.	Population
Suarez-Almazor M, Ortiz Z, Lopez-Olivo M, Moffett M, Pak C, Skidmore B, *et al*. Infliximab and etanercept in rheumatoid arthritis: systematic review of long-term clinical effectiveness, safety, and cost-effectiveness. *Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH)* 2007;**32**.	Population
Summers KM, Kockler DR. Rituximab treatment of refractory rheumatoid arthritis. *Ann Pharmacother* 2005;**39**:2091–5.	Population
Taylor PC. Is abatacept an effective treatment for patients with RA who do not respond to other anti-TNF treatments? Commentary. *Nat Clin Pract Rheumatol* 2006;**2**:128–9.	Design
Van De Putte LBA, Atkins C, Malaise M, Sany J, Russell AS, Van Riel PLCM, *et al*. Efficacy and safety of adalimumab as monotherapy in patients with rheumatoid arthritis for whom previous disease modifying antirheumatic drug treatment has failed. *Ann Rheum Dis* 2004;**63**:508–16.	Population
Van Der Kooij SM, De Vries-Bouwstra JK, Goekoop-Ruiterman YP, Ewals JA, Han KH, Hazes JM, *et al*. Patient-reported outcomes in a randomised trial comparing four different treatment strategies in recent-onset rheumatoid arthritis. *Arthritis Rheum* 2009;**61**:4–12.	Population
Van Der Kooij SM, De Vries-Bouwstra JK, Goekoop-Ruiterman YPM, Ewals JAPM, Han KH, Hazes JMW, *et al*. Patient-reported outcomes in a randomised trial comparing four different treatment strategies in recent-onset rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2009;**61**:4–12.	Population
Van Vollenhoven R, Harju A, Brannemark S, Klareskog L. Treatment with infliximab (Remicade) when etanercept (Enbrel) has failed or vice versa: data from the STURE registry showing that switching tumour necrosis factor alpha blockers can make sense. *Ann Rheum Dis* 2003;**62**:1195–8.	Participant number
Van Vollenhoven RF. Switching between biological agents. *Clin Exp Rheumatol* 2004;**22**:S115–S121.	Design

Article	Reason for exclusion
Van Vollenhoven RF. Switching between anti-tumour necrosis factors: trying to get a handle on a complex issue. *Ann Rheum Dis* 2007;**66**:849–51.	Design
Van VR, Harju A, Brannemark S, Klareskog L. Treatment with infliximab (Remicade) when etanercept (Enbrel) has failed or vice versa: data from the STURE registry showing that switching tumour necrosis factor alpha blockers can make sense. *Ann Rheum Dis* 2003;**62**:1195–8.	Participant number
Venkateshan SP, Sidhu S, Malhotra S, Pandhi P. Efficacy of biologicals in the treatment of rheumatoid arthritis: a meta-analysis. *Pharmacology* 2009;**83**:1–9.	Population
Vera-Llonch M, Massarotti E, Wolfe F, Shadick N, Westhovens R, Sofrygin O, *et al.* Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to tumour necrosis factor-alpha antagonists. *J Rheumatol* 2008;**35**:1745–53.	Design
Villamayor BL, Moreno Ramos MJ, Urbieta SE, Martinez PM, Jorge V, Gonzalez Perez-Crespo C, *et al.* Study of adalimumab's use in rheumatoid arthritis. *Atencion Farmaceutica* 2006;**8**:157–62.	Population
Vital EM, Dass S, Buch MH, Rawstron AC, Ponchel F, McGonagle D, *et al.* Re-treatment of rheumatoid arthritis patients who were initial nonresponders to rituximab: comment on the article by Thurlings *et al. Arthritis Rheum* 2009;**60**:1867.	Design
Voulgari PV, Alamanos Y, Nikas SN, Bougias DV, Temekonidis TI, Drosos AA. Infliximab therapy in established rheumatoid arthritis: an observational study. *Am J Med* 2005;**118**:515–20.	Population
Walsh CAE, Minnock P, Slattery C, Kennedy N, Pang F, Veale DJ, *et al.* Quality of life and economic impact of switching from established infliximab therapy to adalimumab in patients with rheumatoid arthritis. *Rheumatology* 2007;**46**:1148–52.	Population
Weaver AL, Lautzenheiser RL, Schiff MH, Gibofsky A, Perruquet JL, Luetkemeyer J, *et al.* Real-world effectiveness of select biologic and DMARD monotherapy and combination therapy in the treatment of rheumatoid arthritis: results from the RADIUS observational registry. *Curr Med Res Opin* 2006;**22**:185–98.	Population
Weisman MH, Paulus HE, Burch FX, Kivitz AJ, Fierer J, Dunn M, *et al.* A placebo-controlled, randomised, double-blinded study evaluating the safety of etanercept in patients with rheumatoid arthritis and concomitant comorbid diseases. *Rheumatology* 2007;**46**:1122–5.	Population
Witte F. How beneficial is switching from one anti-TNF-alpha agent to a second anti-TNF-alpha agent in patients with rheumatoid arthritis? *Aktuelle Rheumatologie* 2007;**32**:182.	Design
Yazici Y, Yazici H. Tumour necrosis factor alpha inhibitors, methotrexate or both? An inquiry into the formal evidence for when they are to be used in rheumatoid arthritis. *Clin Exp Rheumatol* 2008;**26**:449–52.	Population
Yazici Y, Krasnokutsky S, Barnes JP, Hines PL, Wang J, Rosenblatt L. Changing patterns of tumour necrosis factor inhibitor use in 9074 patients with rheumatoid arthritis. *J Rheumatol* 2009;**36**:907–13.	Population
Yukawa N, Mimori T. [B cell depletion therapy using anti-CD20 antibodies in rheumatoid arthritis.] *Clin Calcium* 2007;**17**:569–76.	Design
Zhang W, Bansback N, Guh D, Li X, Nosyk B, Marra CA, *et al.* Short-term influence of adalimumab on work productivity outcomes in patients with rheumatoid arthritis. *J Rheumatol* 2008;**35**:1729–36.	Population
Zintzaras E, Dahabreh IJ, Giannouli S, Voulgarelis M, Moutsopoulos HM. Infliximab and methotrexate in the treatment of rheumatoid arthritis: a systematic review and meta-analysis of dosage regimens. *Clin Ther* 2008;**30**:1939–55.	Population

Appendix 5

Cost-effectiveness: table of excluded studies with rationale

Article	Reason for exclusion
Bansback N, Ara R, Karnon J, Anis A. Economic evaluations in rheumatoid arthritis: a critical review of measures used to define health states. *Pharmacoeconomics* 2008;**26**:395–408.	Review of clinical measures in rheumatoid arthritis
Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. *Ann Rheum Dis* 2005;**64**:995–1002.	Population
Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. *Health Technol Assess* 2004;**8**(11).	Population
Bullano MF, McNeeley BJ, Yu YF, Quimbo R, Burawski LP, Yu EB, *et al*. Comparison of costs associated with the use of etanercept, infliximab, and adalimumab for the treatment of rheumatoid arthritis. *Manag Care Interface* 2006;**19**:47–53.	Population
Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, *et al*. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. *Health Technol Assess* 2006;**10**(42).	Population
Chiou C-F, Choi J, Reyes CM. Cost-effectiveness analysis of biological treatments for rheumatoid arthritis. *Expert Rev Pharmacoecon Outcomes Res* 2004;**4**:307–15.	Population
Davies A, Cifaldi MA, Segurado OG, Weisman MH. Cost-effectiveness of sequential therapy with tumor necrosis factor antagonists in early rheumatoid arthritis. *J Rheumatol* 2009;**36**:16–25.	Population
Doan QV, Chiou C-F, Dubois RW. Review of eight pharmacoeconomic studies of the value of biologic DMARDs (adalimumab, etanercept, and infliximab) in the management of rheumatoid arthritis. *J Manag Care Pharm* 2006;**12**:555–69.	Review of TNF inhibitors in rheumatoid arthritis
Kamal KM, Miller L-A, Kavookjian J, Madhavan S. Alternative decision analysis modeling in the economic evaluation of tumur necrosis factor inhibitors for rheumatoid arthritis. *Semin Arthritis Rheum* 2006;**36**:50–60.	Review of decision modelling in economic evaluations of TNF inhibitors in rheumatoid arthritis
Kobelt G, Jonsson L, Young A, Eberhardt K. The cost-effectiveness of infliximab (Remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study. *Rheumatology* 2003;**42**:326–35.	Population
Kobelt G, Eberhardt K, Geborek P. TNF inhibitors in the treatment of rheumatoid arthritis in clinical practice: costs and outcomes in a follow up study of patients with Ra treated with etanercept or infliximab in southern Sweden. *Ann Rheum Dis* 2004;**63**:4–10.	Population
Launois R, Payet S, Saidenberg-Kermanac'h N, Francesconi C, Franca LR, Boissier M-C. Budget impact model of rituximab after failure of one or more TNFalpha inhibitor therapies in the treatment of rheumatoid arthritis. *Joint Bone Spine* 2008;**75**:688–95.	Design
Lyseng-Williamson KA, Foster RH. Infliximab: a pharmacoeconomic review of its use in rheumatoid arthritis. *Pharmacoeconomics* 2004;**22**:107–132.	Population
Lyseng-Williamson KA, Plosker GL. Etanercept: a pharmacoeconomic review of its use in rheumatoid arthritis. *Pharmacoeconomics* 2004;**22**:1071–95.	Population
Merkesdal S, Ruof J, Mittendorf T, Zeidler H. Cost-effectiveness of TNF-A-blocking agents in the treatment of rheumatoid arthritis. *Expert Opin Pharmacother* 2004;**5**:1881–6.	Review of TNF inhibitors in rheumatoid arthritis
Monteiro RDC, Zanini AC. Cost analysis of drug therapy in rheumatoid arthritis. *Braz J Pharm Sci* 2008;**44**:25–33.	Population
Muller-Ladner U. Cost effectiveness of biologics in the treatment of rheumatoid arthritis. *Internist* 2004;**45**:1402–6.	Population
Nuijten MJ, Engelfriet P, Duijn K, Bruijn G, Wierz D, Koopmanschap M. A cost-cost study comparing etanercept with infliximab in rheumatoid arthritis. *Pharmacoeconomics* 2001;**19**:1051–64.	Population
Prokes M. Effectiveness of TNF antagonists in routine clinical practice and costs. *Vnitr Lek* 2009;**55**:45–53.	Population
Ravasio R, Lucioni C. Economic evaluation of etanercept in AR. *Pharmacoeconomics – Italian Research Articles* 2006;**8**:129–40.	Review of etanercept

Article	Reason for exclusion
Regier DA, Bansback N, Dar SA, Marra CA. Cost-effectiveness of tumor necrosis factor-alpha antagonist in rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis. *Expert Rev Pharmacoecon Outcomes Res* 2007;**7**:155–69.	Review of TNF inhibitors in rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis
Rubio-Terres C, Ordovas Baines JP, Pla PR, Martinez NC, Sanchez Garre MJ, Rosado Souviron MA. Use and cost of biological disease -modifying anti-rheumatic drugs in Spain (PRAXIS study). *Farm Hosp* 2007;**31**:78–92.	Population
Rubio-Terres C, Ordovas Baines JP, Pla PR. Critical analyis of the article: Use and cost of biological disease-modifying anti-rheumatic drugs in Spain (PRAXIS study). *Farm Hosp* 2008;**32**:190–3.	Population
Suka M, Yoshida K. [Economic evaluation of a new treatment for rheumatoid arthritis.] *Nippon Rinsho* 2007;**65**:1327–30.	Population
Tsutani K, Igarashi A. [Anti-rheumatoid biologics and pharmacoeconomic evaluation.] *Nippon Rinsho* 2005;**63**:711–18.	Design
Unit of Health Economics and Technology Assessment. *Rituximab in patients with rheumatoid arthritis:systematic review and economic evaluation*. Budapest: Unit of Health Economics and Technology Assessment in Health Care (HUNHTA); 2006.	Population
Van Den Hout WB, Goekoop-Ruiterman YPM, Allaart CF, Vries-Bouwstra JKD, Hazes JMM, Kerstens PJSM, *et al*. Cost-utility analysis of treatment strategies in patients with recent-onset rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2009;**61**:291–9.	Population
Virkki LM, Konttinen YT, Peltomaa R, Suontama K, Saario R, Immonen K, *et al*. Cost-effectiveness of infliximab in the treatment of rheumatoid arthritis in clinical practice. *Clin Exp Rheumatol* 2008;**26**:1059–66.	Population
Wailoo AJ, Bansback N, Brennan A, Michaud K, Nixon RM, Wolfe F. Biologic drugs for rheumatoid arthritis in the medicare program: a cost-effectiveness analysis. *Arthritis Rheum* 2008;**58**:939–46.	Population
Walsh CAE, Minnock P, Slattery C, Kennedy N, Pang F, Veale DJ, *et al*. Quality of life and economic impact of switching from established infliximab therapy to adalimumab in patients with rheumatoid arthritis. *Rheumatology* 2007;**46**:1148–52.	Population
Wong JB, Singh G, Kavanaugh A. Estimating the cost-effectiveness of 54 weeks of infliximab for rheumatoid arthritis. *Am J Med* 2002;**113**:400–8.	Population
Wong JB. Cost-effectiveness of anti-tumor necrosis factor agents. *Clin Exp Rheumatol* 2004;**22**:S65–S70.	Review of TNF inhibitors in rheumatoid arthritis
Wu EQ, Chen L, Birnbaum H, Yang E, Cifaldi M. Cost of care for patients with rheumatoid arthritis receiving TNF-antagonist therapy using claims data. *Curr Med Res Opin* 2007;**23**:1749–59.	Population

Appendix 6

Clinical effectiveness: full paper inclusion/exclusion checklist

Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor – full-text inclusion checklist for clinical effectiveness

	Question	Yes	No
Q1	*Population* Did the study include a majority (> 50%) of adults with active rheumatoid arthritis who have had an inadequate response to a TNF inhibitor?	Go to Q2	Exclude UD4 = excluded pop
Q2	*Interventions* Did the interventions include at least one of the following drugs: Adalimumab? Etanercept? Infliximab? Rituximab? Abatacept?	Go to Q3	Exclude UD4 = excluded int
Q3	*Outcomes* Did the study report any clinical outcomes related to efficacy, safety or tolerability?	Go to Q4	Exclude UD4 = excluded out
Q4	*Study design* Was it a primary study (except case reports) or a systematic review?	For primary study: go to Q5 For systematic review: include; UD4 = SR	Exclude UD4 = excluded des
Q5	*Study duration* Was the study at least 12 weeks duration?	Go to Q6	Exclude UD4 = excluded dur
Q6	*Participant numbers* If the study was not an RCT, did it include at least 20 patients in at least one of the treatment arms (if there was more than one arm)?	Include UD4 = included	Exclude UD4 = excluded num

des, design; dur, duration; int, intervention; num, numbers; pop, population; SR, systematic review; UD4, REFERENCE MANAGER User Defined Field 4.

Appendix 7

Cost-effectiveness: full paper inclusion/exclusion checklist

Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor – full-text inclusion checklist for cost-effectiveness

	Question	Yes	No
Q1	*Population* Did the study include a majority of adults with active rheumatoid arthritis who have had an inadequate response to a TNF inhibitor?	Go to Q2	Exclude UD5 = excluded pop
Q2	*Interventions* Did the interventions include at least one of the following drugs: Adalimumab? Etanercept? Infliximab? Rituximab? Abatacept?	Go to Q3	Exclude UD5 = excluded int
Q3	*Outcomes* Did the study report any quality of life estimates, cost estimates or cost-effectiveness results?	Go to Q4	Exclude UD5 = excluded out
Q4	*Study design* Was it a cost–consequence analysis, cost–benefit analysis, cost-effectiveness analysis, cost–utility analysis, cost study (UK only), or quality of life study?	Include UD5 = included	Exclude UD5 = excluded des

des, design; dur, duration; int, intervention; num, numbers; pop, population; UD5, REFERENCE MANAGER User Defined Field 5.

Appendix 8

Clinical effectiveness review: data extraction form

Adalimumab/Etanercept/Infliximab/Rituximab/Abatacept (delete as appropriate)

RCT/Controlled study (concurrent)/Controlled study (historical)/Uncontrolled study (delete as appropriate)

First author and *year*		Reference no.	
Trial name/protocol no.		Reviewer	
Citation		Date of abstraction	
Country and no. of centres		Sponsorship	
Related references			

Inclusion criteria	**General comments and comments on exclusions**
Age:	
Duration of RA ≥	
Prior TNF inhibitor treatment: Reason for discontinuation of TNF inhibitor:	
Disease activity parameters	
Tender joint count ≥	
Swollen joint count ≥	
ESR ≥	
CRP ≥	
Morning stiffness >	

Other inclusion/exclusion criteria:
Concomitant treatments during the trial Methotrexate: allowed/not allowed/unclear/conditional: Other DMARDs: allowed/not allowed/unclear/conditional: Steroids: allowed/not allowed/unclear/conditional: Other treatments allowed: Other treatments not allowed:
Previous TNF inhibitor(s) Eligibility for the previous anti-TNF: Doses and treatment duration of previous TNF inhibitor (and concomitant DMARDs): Washout period from the previous TNF inhibitor:

Randomised controlled trial study design and quality

Was randomisation adequate: Yes/No/Unclear
Was allocation adequately concealed: Yes/No/Unclear
Blinding:
Were patients blinded from the study interventions: Yes/No/Unclear
Were study investigators/outcome assessors blinded from the study interventions: Yes/No/Unclear
Were data analysts blinded from the study interventions: Yes/No/Unclear
Was lost to follow-up stated for each treatment groups: Yes/No/Unclear
Was ITT analysis used: Yes/No/Unclear
Duration of treatment: Duration of follow-up (if different):
Study visits (outcome data available):
Comments on study design and quality (problem in study design; power of study; potential bias):

Non-randomised controlled trial study design and quality

What was the study design:
Were criteria for including patients into the study stated?
Were consecutive patients meeting the inclusion criteria (if any) entered into the study?
Was lost to follow-up stated for each treatment groups: Yes/No/Unclear
Duration of treatment: Duration of follow-up (if different):
Study visits (outcome data available):
Comments on study design and quality (problem in study design; power of study; potential bias):

Interventions and comparators

State drug name(s), dose, frequency, route of administration
A)
B)
C)
D)
E)
F)

Baseline characteristics

Tx arm	A)	B)	C)	D)	E)	F)	All patients
Patient number							
Age (mean, yrs)							
Female %							
Disease duration (yrs)							
Auto antibody status							
(Comorbidity) %							
(Comorbidity) %							
(Comorbidity) %							
Previous TNF inhibitor							
No. of previous DMARDs							
(Previous DMARD) %							
(Previous DMARD) %							
On steroids (%)							
On NSAIDs (%)							
If on MTX – dose?							
% joint replm							
Comments on the presence or absence of significant differences between treatment arms: No. of patients screened: No. of patients randomised: No. of patients received at least one dose of study drug:							

Outcomes

ITT population/efficacy population (delete as appropriate)

Measure of activity	***Values (SD or IQR)***	***Intervention – A n=***	***Intervention – B n=***	***Intervention – C n=***	***Intervention – D n=***	***Intervention – E n=***	***Intervention – F n=***
1. Withdrawal – lack of efficacy	*No. eval.* *No. withdrew*						
2. Withdrawal –adverse events	*No. eval.* *No. withdrew*						
3. Withdrawal – any reason	*No. eval.* *No. withdrew*						
4. ACR20 %	*No. eval.* *No. improved*						
5. ACR50 %	*No. eval.* *No. improved*						
6. ACR70 %	*No. eval.* *No. improved*						
7. Swollen joint count ( ) *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						

8. Tender joint count ( ) *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						
9. Pain – patient ( ) *Specify week*	*No. eval.*						
	Pre-Rx						
Post							
	Chge *p-value*						
10. Phys. Global ( ) *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						
11. Patient global ( ) *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						
Measure of activity	***Values (SD or IQR)***	***Intervention – A n=***	***Intervention – B n=***	***Intervention – C n=***	***Intervention – D n=***	***Intervention – E n=***	***Intervention – F n=***
12. CRP *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						
13. ESR *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						
14. HAQ *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						
15. DAS *Specify week*	*No. eval.*						
	Pre-Rx						
	Post						
	Chge *p-value*						
16. Joint damage	*No. eval.*						
(scale:	*Pre-Rx*						
	Post						
Specify week	*Chge* *p-value*						

(scale:	*Pre-Rx*						
	Post						
Specify week	*Chge* *p-value*						
Comments: Which is/are the primary end point(s)? How were missing data handled (e.g. LOCF)? Were any outcome evaluation planned but not reported? Results of subgroup analysis							

Adverse events

	Interventions:					
	A **(*n*=)**	**B** **(*n*=)**	**C** **(*n*=)**	**D** **(*n*=)**	**E** **(*n*=)**	**F** **(*n*=)**
Deaths						
Serious adverse events						
Serious infection (definition:						
Infections needing antibiotics						
Any infection						
Malignancy						
Injection site reaction						
Infusion reaction						
Others:						
Comments:						

Appendix 9

Cost-effectiveness review: data extraction of included studies

TABLE 108 Lindgren *et al.*:[145] economic evaluation data extraction

Author	Lindgren	**Date**	2009	**Study population**	Patients with active RA and an inadequate response to one or more TNF inhibitor agents	**Type of economic evaluation**	Cost–utility analysis
Intervention	RTX						
Clinical effectiveness							
Source of effectiveness data	Effectiveness of treatment with TNF inhibitors is based on patient level data from the Southern Swedish Arthritis Treatment Group Registry (SSATG) (1997–2007) This data set contains baseline demographic data, disease information (all available HAQ and DAS28 scores), treatment data (biologics and DMARDs) and utility scores EQ-5D The data set used for this analysis contained 1,903 patients with sufficient data on up to three lines of treatment Source for RTX effectiveness was the REFLEX trial,[124–126] where patients with active RA and an inadequate response to one or more TNF inhibitors were randomised to receive i.v. RTX (one course, two infusions of 1,000 mg each) or placebo, both with MTX as background therapy			**Clinical outcomes measured and methods of valuation used**	REFLEX primary efficacy point was ACR20 response at 6 months. Secondary end points were ACR50 and ACR70 response, DAS28, and EULAR response criteria at 6 months Mean HAQ scores declined from 1.9 to 1.4 at the 4-week measurement and remained constant up to 6 months of treatment Mean DAS28 scores declined from 6.9 to 5.4 after 4 weeks and to 5.0 after 6 months. Assuming normal distribution of the scores, 5.9% of patients would achieve a DAS28 below 3.2 at week 4, but no further change to low disease activity thereafter Utilities are mapped from the HAQ score. The model uses the equation as estimated by SSATG data (6,860 observations for 1,787 patients) Quality of life (QoL) = 0.915 – 0.252 × HAQ – 0.05 × male – 0.107 × DAS28 HAQ progression was estimated through the SSATG data. It is unclear though what type of regression was used; text suggests linear while Table 2 suggests logistic. Also, Table 2 should have a clearer indication of which variable is the dependent one on all functions used HAQ progression = 0.106 + 0.241 × (HAQ at treatment start) + 0.002 × (months on treatment) – 0.087 × (second line) – 0.192 × (third line) 0.007 × (disease duration)		

continued

TABLE 108 Lindgren *et al*.:[145] economic evaluation data extraction *(continued)*

Cost data							
Currency used	Costs estimated in Swedish krona (SEK) and presented in euro (€) (1€ = 9.45 SEK)	**Years to which costs apply**	2008	**Perspective(s)**	Societal perspective (direct and indirect costs included as well as informal care)		
Cost data handled appropriately	Yes Source for resource consumption was a survey carried out at regular intervals by the department of rheumatology at the Malmö University Hospital (southern Sweden). The survey covers an estimated 90% of the patient population in the area and includes all costs, direct medical and non-medical, as well as productivity losses Costs were calculated as a function of HAQ and DAS28 The cost of TNF inhibitor treatment was a weighted mean based on usage of each drug Unit costs were obtained from standard national (Swedish) sources The cost of RTX was based on the dose used in REFLEX[124–126] (two infusions of 1,000 mg each per course). Retreatment could take place between 4 and 12 months, at a 6-month interval Costs of AEs (such as hospitalisation due to severe infections or clinical investigations) were excluded from the analysis as such costs would occur in both arms Costs are discounted at 3%						
Cost-effectiveness							
Modelling summary	A discrete event simulation model was developed Patients in the model can be in three states: on treatment, off treatment or dead. On treatment, a difference is made between the first, second or third TNF inhibitors, but not between the different agents. The treatment state is further divided into high or low disease activity, with the cut-off point defined as DAS28 = 3.2 Simulation starts when patients start on second-line treatment, either with a second TNF inhibitor or with RTX. Patients will stay on these treatments until discontinuation of the second-line TNF inhibitor (according to SSATG data) or withdrawal from RTX (according to data from REFLEX[124–126]). Patients previously on RTX will receive their second TNF inhibitor. When patients fail again, they will switch to another TNF inhibitor again. In the absence of sufficient data to estimate the event rates for the fourth (or subsequent) TNF treatment lines, these are assumed to be the same as for the third line Improvement in HAQ score was assumed to occur immediately and HAQ levels thereafter were assessed using linear regression (as indicated in text – not clear on the table) on the difference compared with the initial HAQ response. At treatment discontinuation, patients return to the initial HAQ score and progress at the rate of 0.03 per year while off treatment Base case is for a 52-year-old female patient with a HAQ score of 1.9 at the start of the second biologic and disease duration of 12 years						
Outcome measures used in economic evaluations	Incremental QALYs and ICERs	**Statistical analysis for patient-level stochastic data**	A Cox proportional hazards model was estimated to identify covariates (age, gender, disease duration, current HAQ, current disease activity, treatment line) with a possible impact on times to event Bootstrapping was used for parameters where patient-level data were available	**Appropriateness of statistical analysis**	Yes		

TABLE 108 Lindgren *et al*.:[145] economic evaluation data extraction *(continued)*

Uncertainty			
Uncertainty around cost-effectiveness expressed	Yes Model uncertainty was explored using PSA with 1,000 samples by Monte Carlo simulation using all available data and patient characteristics	**Appropriateness of method dealing with uncertainty around cost-effectiveness**	Yes
Sensitivity analysis	Sensitivity analysis for the key variables was performed For parameters relating to RTX and the progression of HAQ, normal distribution was assumed	**Modelling inputs and techniques appropriate**	Yes
Author's conclusions	The strategy including RTX in second line dominates current treatment Total costs were €401,000 for the RTX arm and €403,600 for current treatment Patients in the RTX arm gain 0.20 additional QALYs, owing in part to the absence of lag-time in restarting a TNF inhibitor at withdrawal of RTX Changes in the individual key parameters do not affect these results Only if RTX was administered every 4 months or less, then costs for this strategy are higher The results from the PSA indicate that all but one of the 1,000 simulations fall below a theoretical threshold of 500,000 SEK (€53,000)		

TABLE 109 Russell *et al.*:[143] economic evaluation data extraction

Author	Russell	**Date**	2009	**Study population**	Patients with moderate to severe RA and with an inadequate response to one or more DMARDs and/or TNF inhibitors	**Type of economic evaluation**	Cost-effectiveness analysis
Intervention	ABT						
Clinical effectiveness							
Source of effectiveness data	DAS data are from various published sources, including the ATTAIN[127–132] and TEMPO trials The AIM trial was the source for patients' inadequate response to DMARDs; safety and effectiveness of ABT when appearing in the sequence for the first time (TNF inhibitor inadequate responders); effectiveness of ABT maintained after the first cycle and for one or more subsequent 6-month cycles The ATTAIN trial[127–132] was the source for patients' inadequate response to TNF inhibitor therapies; safety of ABT, effectiveness of ABT maintained after the first cycle and for one or more subsequent 6-month cycles The TEMPO trial was the source for effectiveness of ETN when appearing in the sequence for the first time (DMARD inadequate responders); effectiveness of ETN maintained after the first cycle and for one or more subsequent 6-month cycles			Clinical outcomes measured and methods of valuation used	Treatment effectiveness was defined as either achieving disease remission (DAS28 < 2.6) or low disease activity rate (DAS28 ≤ 3.2) The effectiveness of TNF inhibitors in TNF inhibitor inadequate responders was extracted from the ATTAIN trial,[127–132] assuming a 10% reduction after each switch		
Cost data							
Currency used	CAN$	**Years to which costs apply**	2006	**Perspective(s)**	Public payer		
Cost data handled appropriately	ABT is administered over a 30-minute i.v. infusion at 2 and 4 weeks after the first infusion, and every 4 weeks thereafter The analysis assumes an average dose of 750 mg (3 × 250 mg vials) per infusion However, infusion costs were not included because in Canada, IFX and ABT were administered in participating rheumatology and infusion clinics or at home for ABT Direct medical costs per DAS score categories were assessed based on a Canadian cost survey. Data were collected from 253 adult patients and the following cost categories were collected: visits to health professionals [family physician, specialist (non-surgical reported separately from surgical visits), allied health, dentist], laboratory tests or investigation (X-ray, computerised tomography, magnetic resonance imaging, ultrasound, electrocardiogram, other laboratory, bone density), hospitalisations, prescribed drugs [arthritis (not including TNF inhibitor or costimulation modulator), antihypertensive, gastroprotective, other], home care, transportation services, adaptive aids/other devices The estimated annual costs of therapy were: ABT (250-mg vial): $18,480 (year 1), $17,160 (year 2) ADA (40-mg pre-filled syringe): $17,680 (year 1), $17,680 (year 2) ETN (25-mg vial): $18,200 (year 1), $18,200 (year 2) IFX (100-mg vial): $20,445 (year 1), $18,330 (year 2)						
Cost-effectiveness							
Modelling summary	Fourteen decision trees (for the various strategies) were designed and analysed as simulation models in DECISIONPRO software Patients with moderate-to-severe RA with an inadequate response to DMARDs, eligible for biologic therapy are entering the model						

TABLE 109 Russell *et al.*:[143] economic evaluation data extraction *(continued)*

	Patients achieving treatment success (defined as achieving either a low disease activity rate or remission) are maintained on existing therapy for up to 2 years. Those with an inadequate response to a biologic therapy are switched to a subsequent biologic agent, with decision to switch made at 6-month intervals in case of an inadequate response The model assesses the cost-effectiveness of ABT used as first biologic therapy in patients with an inadequate response to DMARDs and as second biologic therapy in patients with an inadequate response to a first TNF inhibitor The comparator was defined as a successive trial of TNF inhibitor therapies based on the most established treatment pattern in Canada at time of model development. RTX was not reimbursed for RA in Canada at that time, therefore it was not considered as a valid comparator The same treatment continues as long as it is efficacious; decision to switch treatment for all causes (lack or loss of efficacy, adverse events, intolerance, etc.); the model allows switches to occur every 6 months The model calculates the overall effectiveness of each entire sequence of biologic strategies as an effectiveness outcome Reference case was a 2-year treatment with up to three successive biologic agents (in case of an inadequate response to the previous biologic agent) ETN → IFX → ADA → DMARDs The following strategies were simulated: ABT → ETN → IFX → DMARDs ETN → ABT → IFX → DMARDs				
Outcome measures used in economic evaluations	Cost per additional case of LDAS gained Cost per additional remission gained	**Statistical analysis for patient-level stochastic data**	Not undertaken	**Appropriateness of statistical analysis**	NA
Uncertainty					
Uncertainty around cost-effectiveness expressed	PSA using 5,000 Monte Carlo simulations was used to explore uncertainty in the model Beta distribution was used for transition probabilities; log-normal distribution was used for costing variability	**Appropriateness of method dealing with uncertainty around cost-effectiveness**	Yes		
Sensitivity analysis	One-way sensitivity analyses (scenario based) was undertaken	**Modelling inputs and techniques appropriate**	Yes		
Author's conclusions	*Inadequate response to DMARDs – cost per additional case of LDAS gained* The lowest cost biologic strategy was ABT used as the first biologic agent. This strategy dominated the other two, providing 13.8% greater probability (29.4% vs 15.6%) of achieving LDAS than sequential TNF inhibitor therapy with an overall RA-related cost saving of $730 ($39,759 vs $40,489) over 2 years ABT used as a second biologic after an inadequate response to one TNF inhibitor (ETN) was cost-effective, providing 3.7% greater probability of achieving LDAS (19.3% vs 15.6%) at an additional cost of $463 ($40,952 vs $40,489) over the 2-year period, with an ICER of $12,514 per additional case of LDAS gained Thus, ABT used as first biologic appears to be less costly and to provide greater probability of achieving LDAS than using ABT as second biologic agent *Inadequate response to DMARDs – cost per additional remission gained* The lowest cost biologic strategy was ABT used as the first biologic agent. This strategy dominated the other two, providing 9.6% greater probability (14.8% vs 5.2%) of remission than sequential TNF inhibitor therapy with an overall RA-related cost-saving of $504 ($38,061 vs $38,565) over 2 years ABT used as a second biologic after an inadequate response to one TNF inhibitor (ETN) was cost-effective, providing 3.5% greater probability of achieving remission (8.7% vs 5.2%) at an additional cost of $589 ($39,154 vs $38,565) over the 2-year period, with an ICER of $16,829 per additional remission gained Thus, ABT used as first biologic appears to be less costly and to provide greater probability of achieving remission than using ABT as second biologic agent *Inadequate response to ETN* After an initial 6-months treatment failure to ETN, all patients were switched to either ABT or IFX as the second biologic option, followed by IFX and ADA, respectively ABT used as second biologic agent was cost-effective, providing 6.9% additional treatment success rate for achieving LDAS (17.1% vs 10.2%) and 3.5% additional treatment success rates for achieving remission (7.4% vs 3.9%) at an ICER of £20,377 per additional case of LDAS and $26,400 per additional remission, respectively				

LDAS, low disease activity (DAS ≤ 3.2).

TABLE 110 Kielhorn *et al*.:[144] economic evaluation data extraction

Author	Kielhorn	**Date**	2008	**Study population**	Patients with RA who failed to respond adequately to two non-biologic DMARDs and one TNF inhibitor	**Type of economic evaluation**	Cost–utility analysis
Intervention	RTX						
Clinical effectiveness							
Source of effectiveness data	The mean drop in HAQ for each of the responder groups is calculated from the REFLEX trial[124–126]			**Clinical outcomes measured and methods of valuation used**	Utilities are mapped from the HAQ score. The model uses the equation as estimated by Bansback *et al.*[146] $QoL = 0.76 - 0.28 \times HAQ + 0.05 \times female$ All-cause mortality is derived by GAD[218] and adjusted with an RA risk multiplier related to each individual's HAQ score (Barton *et al.*[175])		
Cost data							
Currency used	British £	**Years to which costs apply**	2004 (not explicitly stated)	**Perspective(s)**	NHS and Personal Social Services		
Cost data handled appropriately	For each treatment, drug cost, administration cost and monitoring cost were considered Drug costs were obtained from *British National Formulary* 50 Administration costs are generated by biological DMARDs requiring infusion or injection For RTX, 5 hours of administration was assumed on average, including pre-medication For IFX, a 3-hour infusion time for the 225 mg of active substance was assumed including post-infusion observation time A weight of 78 kg was assumed (Cohen *et al.*[126]) No drug wastage or increase in dose was included in the calculation Healthcare personnel attendance time was estimated according to Nuijten *et al.*,[219] and personnel salaries were obtained from Personal Social Services Research Unit (PSSRU) 2004 Monitoring costs include an outpatient visit or a GP visit, and certain examination and tests. Costs for these were obtained from NHS, PSSRU or Barton *et al.*[175] Costs are linked to functional status, as measured by the HAQ score, by grouping HAQ scores into six categories (Kobelt *et al.* 1999,[220] 2004[221]). Each HAQ score category was assigned an average cost. Direct costs included the cost of the drug, drug administration, medical resource consumption (comedication, surgery, etc.) All costs accruing after the first year of the evaluation were discounted at 3.5%						

TABLE 110 Kielhorn *et al.*:[144] economic evaluation data extraction *(continued)*

Cost-effectiveness					
Modelling summary	A microsimulation Markov model was designed and analysed in Microsoft EXCEL. A cycle length of 6 months was used. Patients either follow the current standard treatment sequence reflecting real-life clinical practice in the UK or an alternative sequence, which is identical, except for the introduction of RTX as an additional treatment within the sequence. If patients respond they remain on the drug for a predetermined period of time. If they do not respond, they continue to the next treatment in the sequence. They remain in palliative care (MTX) until they reach 100 years of age or death Analysis A assumes non-sequential use of bDMARDs (NICE 36[222]) Analysis B assumed sequential use of bDMARDs; (based on data from the BSRBR and Hyrich *et al.*[223]) Patients enter the model and are allocated to either of the two treatment sequences. The patients are then exposed to the first treatment in the sequence and are allocated to one of the three responder groups ACR 20–49, 50–69, 70 + or to the non-responder group The HAQ score is assumed to drop by 0.1 for non-respondents, 0.45 for ACR20–49, 0.85 for ACR50–69 and 1.11 for ACR70 + respondents (Kielhorn *et al.*[224]). While on treatment, patient HAQ scores are assumed to progress by 0.017 during each cycle of the model (Scott *et al.*[225]). HAQ progression for patients on palliative care is assumed to be 0.065 (Bansback *et al.*[146]) Time on treatment in the sequence was derived from Barton *et al.*,[175] and was 4.25 years for all bDMARDs apart from IFX, for which, driven by a higher drop-out of patients, 2.46 years was assumed. bDMARDs treatment duration was 1.7 years for ciclosporin A, 3.85 years for gold and 4.1 years for LEF. For RTX a course of 2 × 1,000 mg every 9 months over the course of 4.25 years was assumed. For all other drugs the licensed dose as per the EU label was assumed Once treatment stops, the entire initial gain in HAQ score is assumed to be lost instantly (100% rebound effect). Patients are then allocated to the next available treatment option until the treatment sequence is exhausted. At this point, all patients receive palliative care, defined as single agent MTX, until death Patients leave the model when they reach the age of 100 years or die				
Outcome measures used in economic evaluations	Incremental QALYs and ICERs	**Statistical analysis for patient-level stochastic data**	Not undertaken	**Appropriateness of statistical analysis**	NA
Uncertainty					
Uncertainty around cost-effectiveness expressed	Yes Model uncertainty was explored using PSA with 1,000 samples by Monte Carlo simulation. Owing to lack of data it was not possible to run a PSA on all variables. For these variables, one-way sensitivity analysis was applied instead A Dirichlet distribution was fit for response rate parameters, a Weibull distribution into the time on treatment parameters and a normal distribution was fit into the inpatient days, trimmed for values $[0, +\infty)$	**Appropriateness of method dealing with uncertainty around cost-effectiveness**	Yes		
Sensitivity analysis	Yes One-way sensitivity analysis was applied to determine the relative importance of different parameters to the primary outcome The model was not sensitive with respect to changes to assumed time on treatment, or changes between adjusted and unadjusted response rates Larger variability was observed in changes to RTX dosing retreatment from 9 months to 6 months and when changing the HAQ long-term progression Variability was also observed when baseline age is increased	**Modelling inputs and techniques appropriate**	Yes		

continued

TABLE 110 Kielhorn *et al*.:[144] economic evaluation data extraction *(continued)*

Author's conclusions	Both analyses showed higher treatment cost in the sequence containing RTX Analysis A Total discounted QALYs were 3.051 and 2.324 for the RTX arm and the standard of care arm, respectively, resulting in a QALY gain of 0.727. The ICER based on total direct medical costs was £14,690 Analysis B QALY gain was 0.526; the ICER based on total direct medical costs was £11,601

TABLE 111 Vera-Llonch *et al*.:[142] economic evaluation data extraction

Author	Vera-Llonch	**Date**	2008	**Study population**	Women with moderate-to-severe RA with inadequate response to TNF inhibitors	**Type of economic evaluation**	Cost–utility analysis
Intervention	ABT						
Clinical effectiveness							
Source of effectiveness data	Source for effectiveness data was the ATTAIN trial[127–132]			**Clinical outcomes measured and methods of valuation used**	Improvement in HAQ scores during the first 6 months of therapy For patients continuing to receive ABT beyond 6 months, the improvement at 6 months was assumed to persist over time. For patients discontinuing ABT, the HAQ score was assumed to return to a value equal to what it would have been in the absence of such treatment (oral DMARD only) Initial HAQ scores are randomly assigned to each patient entering the model from an assumed initial probability distribution. Future values of the HAQ score were estimated based on the assumed initial value, the expected rate of disease progression and the expected effect of treatment The estimated mean percentage HAQ change 3 months after therapy initiation in ATTAIN[127–132] was 21%; at 6 months it was 25.5% The distribution of the HAQ change with ABT was assumed to be truncated normal, based on visual inspection of the data in ATTAIN[127–132] Among patients continuing to receive ABT, the percentage reduction in the HAQ score was assumed to remain constant at the level prevailing at 6 months. However, the HAQ value against which this percentage reduction was applied was increased by 0.015 annually Health-state utility values were mapped from the HAQ score. Although mean utilities corresponding to the appropriate HAQ score are presented in a table, the exact formula that was using for this mapping is not provided For patients receiving oral DMARD only, the HAQ score was assumed to increase by 0.065 annually to reflect disease progression Mortality risk was estimated through age and the expected value of the HAQ score Health-state utilities were similarly estimated based on the expected future values of the HAQ score		

TABLE 111 Vera-Llonch *et al*.:[142] economic evaluation data extraction *(continued)*

Cost data					
Currency used	US$	**Years to which costs apply**	2006	**Perspective(s)**	Third-party payer (medical treatment only – direct non-medical costs or loss productivity were excluded)
Cost data handled appropriately	Following an initial infusion, ABT was assumed to be administered on days 14 and 29, and every 4 weeks thereafter, for a total of 15 infusions during the first year and 13 infusions every year thereafter Patients weighing < 60 kg were assumed to receive two vials (500 mg) per infusion; 60–100 kg, three vials (750 mg); and > 100 kg, four vials (1 g) The cost of ABT was assumed to be $450 per 250-mg vial The cost of each 30-minute infusion was assumed to be $129 Oral DMARD therapy was assumed to consist of MTX. The annual cost of treatment with MTX was assumed to be $600, based on an assumed dose of 15 mg weekly Estimates of the cost of baseline and routine monitoring for patients receiving ABT were based on product labelling, published guidelines and Medicare payment rates Tests for ABT patients were assumed to cost $9 (one-off cost) while tests for the DMARD patients were at $181 per year Costs were discounted at 3%				
Cost-effectiveness					
Modelling summary	A simulation model of a hypothetical cohort of 1,000 women aged 55–64 years was developed. The model cycle was 3 months Patients enter the model at either the 'oral DMARD' state or the 'oral DMARD state plus ABT' Patients on ABT are assumed to initiate treatment on day 1 [500–1,000 mg (based on body weight) i.v. infusion over 30 minutes], and receive additional infusions on day 14, day 29 and every 4 weeks thereafter Patients with HAQDI improvements of –0.50 or greater at 6 months were assumed to continue to receive ABT Patients failing to achieve this improvement are assumed to discontinue treatment Patients also discontinue treatment for other reasons such as side effects, intercurrent illness and surgery All patients discontinuing ABT are assumed to continue to receive 'oral DMARDs' Authors justify this assumption (assuming no switch from ABT to another biologic DMARD) on the bases that there are no data on the efficacy of the latter agents given prior failure with ABT Time horizons were 10 years and lifetime				
Outcome measures used in economic evaluations	Incremental cost per QALY	**Statistical analysis for patient-level stochastic data**	Not undertaken	**Appropriateness of statistical analysis**	NA
Uncertainty					
Uncertainty around cost-effectiveness expressed	Expressed through 100 Monte Carlo simulations	**Appropriateness of method dealing with uncertainty around cost-effectiveness**	Yes		
Sensitivity analysis	Yes Selected assumptions and parameter estimates were varied, including: Discontinuation of ABT therapy for lack of efficacy or other reasons Timing of therapy discontinuation due to lack of efficacy (3 months vs 6 months) Odds ratio for mortality associated with each 1-point increase in the HAQ score Assumption of mortality benefit with ABT Expected rate of disease progression Threshold for clinical meaningful improvement in HAQ Women aged other than 55–64 years Male population	**Modelling inputs and techniques appropriate**	Yes		

continued

TABLE 111 Vera-Llonch *et al*.:[142] economic evaluation data extraction *(continued)*

Author's conclusions	Over a 10-year time horizon, the cost-effectiveness of ABT was estimated to be $50,576 per QALY gained
	On a lifetime basis, cost-effectiveness was $45,979 per QALY gained
	At a threshold of $100,000 per QALY, the probability that ABT would be cost-effective was 1
	At a threshold of $20,000 per QALY, ABT would be unlikely to be cost-effective (probability = 0)
	At a threshold of $50,000 per QALY, the probability that ABT would be cost-effective was 0.39 over a 10-year time horizon and 1 over lifetime

Appendix 10

Outcomes not reported in the main text of the report

Adalimumab

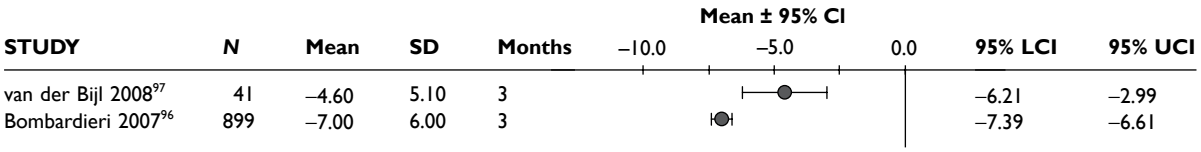


FIGURE 100 Adalimumab: SJC, change from baseline. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

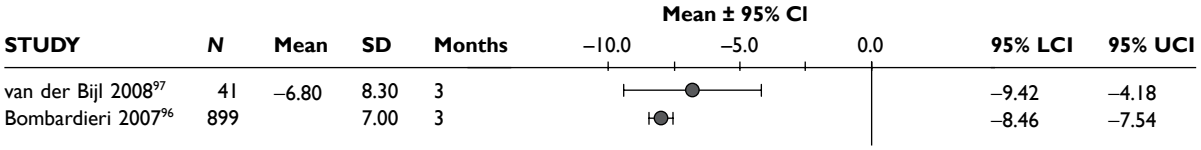


FIGURE 101 Adalimumab: TJC, change from baseline. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

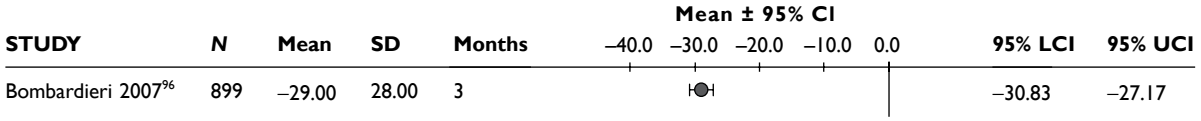


FIGURE 102 Adalimumab: pain (visual analogue scale), change from baseline. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

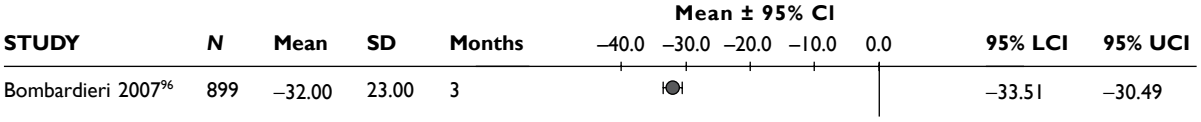


FIGURE 103 Adalimumab: physician global assessment (visual analogue scale), change from baseline. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

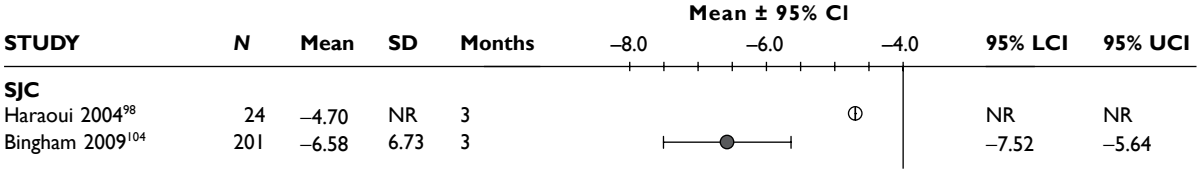
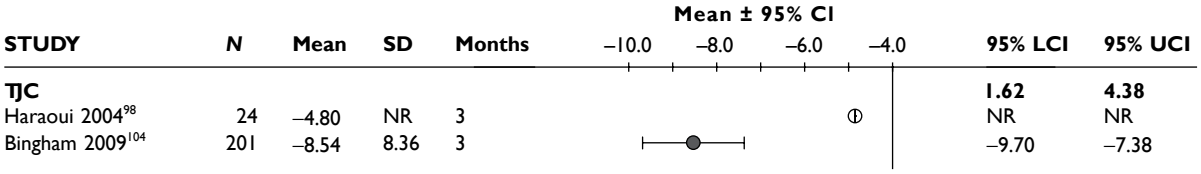
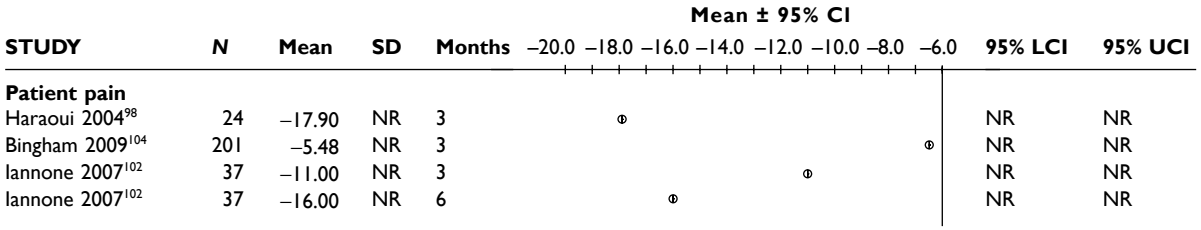


FIGURE 104 Etanercept: SJC. LCI, lower confidence interval; NR, not reported; SD, standard deviation; UCI, upper confidence interval.



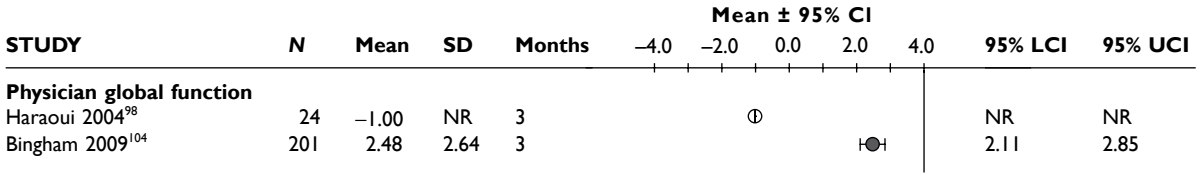
STUDY	N	Mean	SD	Months	95% LCI	95% UCI
TJC					**1.62**	**4.38**
Haraoui 2004[98]	24	−4.80	NR	3	NR	NR
Bingham 2009[104]	201	−8.54	8.36	3	−9.70	−7.38

FIGURE 105 Etanercept: TJC. LCI, lower confidence interval; NR, not reported; SD, standard deviation; UCI, upper confidence interval.



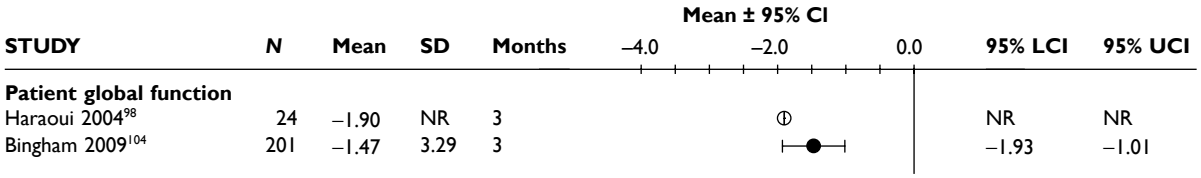
STUDY	N	Mean	SD	Months	95% LCI	95% UCI
Patient pain						
Haraoui 2004[98]	24	−17.90	NR	3	NR	NR
Bingham 2009[104]	201	−5.48	NR	3	NR	NR
Iannone 2007[102]	37	−11.00	NR	3	NR	NR
Iannone 2007[102]	37	−16.00	NR	6	NR	NR

FIGURE 106 Etanercept: patient pain. LCI, lower confidence interval; NR, not reported; SD, standard deviation; UCI, upper confidence interval.



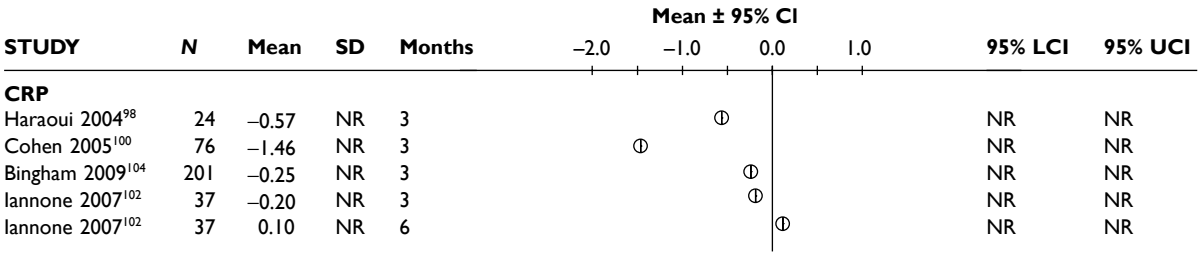
STUDY	N	Mean	SD	Months	95% LCI	95% UCI
Physician global function						
Haraoui 2004[98]	24	−1.00	NR	3	NR	NR
Bingham 2009[104]	201	2.48	2.64	3	2.11	2.85

FIGURE 107 Etanercept: physician global function. LCI, lower confidence interval; NR, not reported; UCI, upper confidence interval.



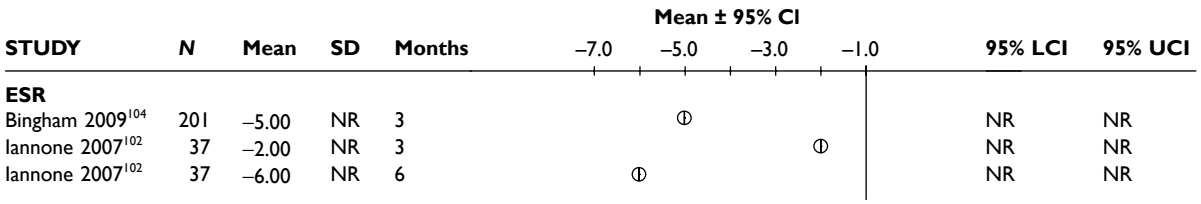
STUDY	N	Mean	SD	Months	95% LCI	95% UCI
Patient global function						
Haraoui 2004[98]	24	−1.90	NR	3	NR	NR
Bingham 2009[104]	201	−1.47	3.29	3	−1.93	−1.01

FIGURE 108 Etanercept: patient global function. LCI, lower confidence interval; NR, not reported; SD, standard deviation; UCI, upper confidence interval.



STUDY	N	Mean	SD	Months	95% LCI	95% UCI
CRP						
Haraoui 2004[98]	24	−0.57	NR	3	NR	NR
Cohen 2005[100]	76	−1.46	NR	3	NR	NR
Bingham 2009[104]	201	−0.25	NR	3	NR	NR
Iannone 2007[102]	37	−0.20	NR	3	NR	NR
Iannone 2007[102]	37	0.10	NR	6	NR	NR

FIGURE 109 Etanercept: mean change from baseline in CRP. LCI, lower confidence interval; NR, not reported; SD, standard deviation; UCI, upper confidence interval.



STUDY	N	Mean	SD	Months	95% LCI	95% UCI
ESR						
Bingham 2009[104]	201	−5.00	NR	3	NR	NR
Iannone 2007[102]	37	−2.00	NR	3	NR	NR
Iannone 2007[102]	37	−6.00	NR	6	NR	NR

FIGURE 110 Etanercept: mean change from baseline in ESR. LCI, lower confidence interval; NR, not reported; SD, standard deviation; UCI, upper confidence interval.

Etanercept

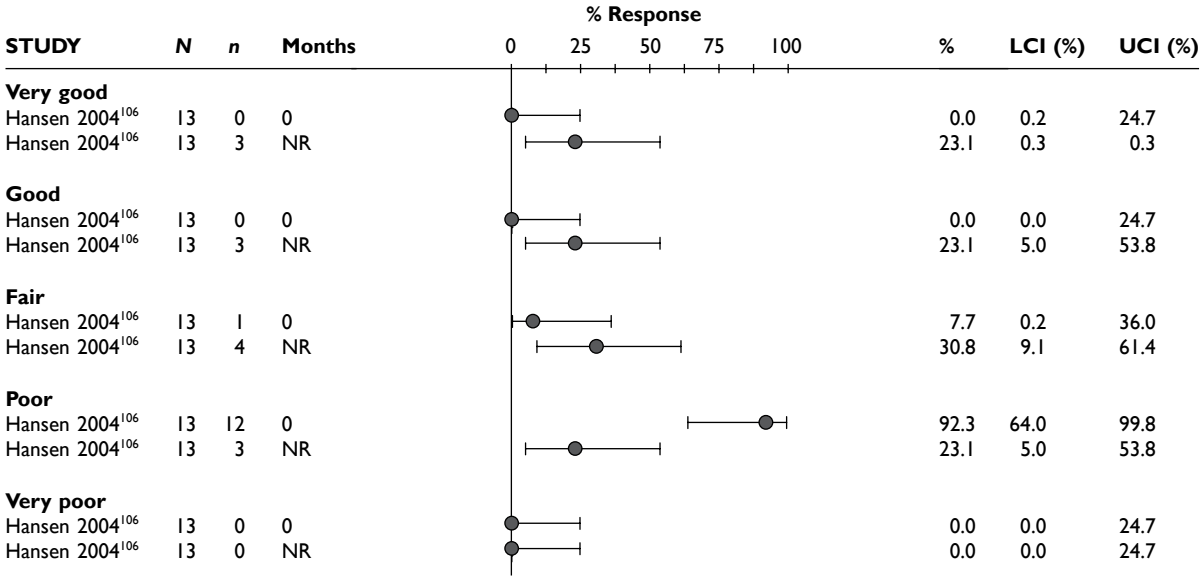


FIGURE 111 Infliximab: physician global assessment. LCI, lower confidence interval; NR, not reported; UCI, upper confidence interval.

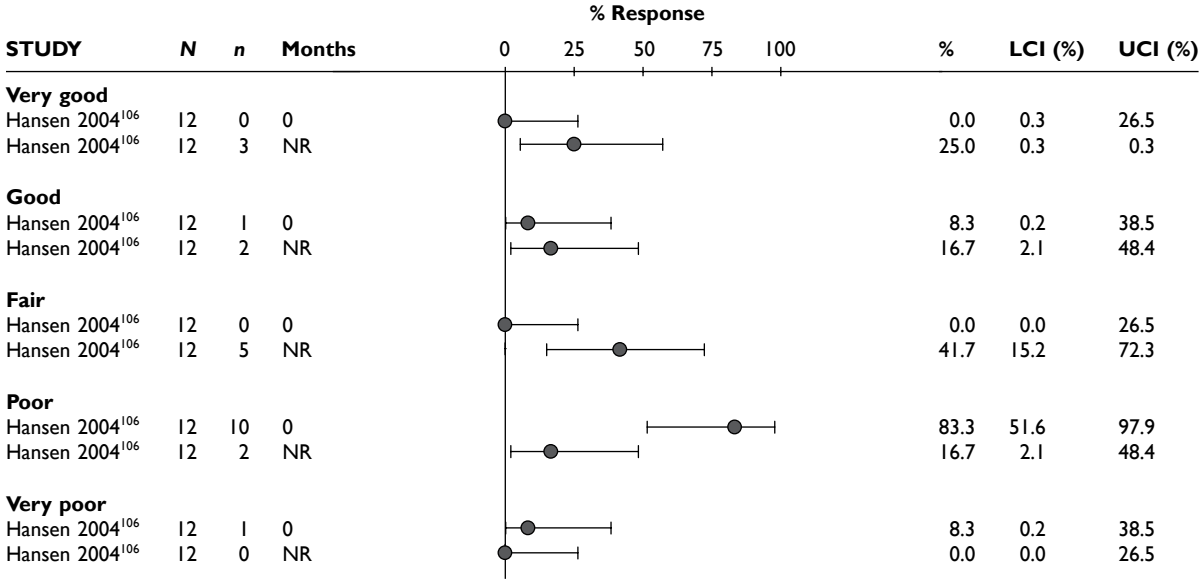


FIGURE 112 Infliximab: patient global assessment. LCI, lower confidence interval; NR, not reported; UCI, upper confidence interval.

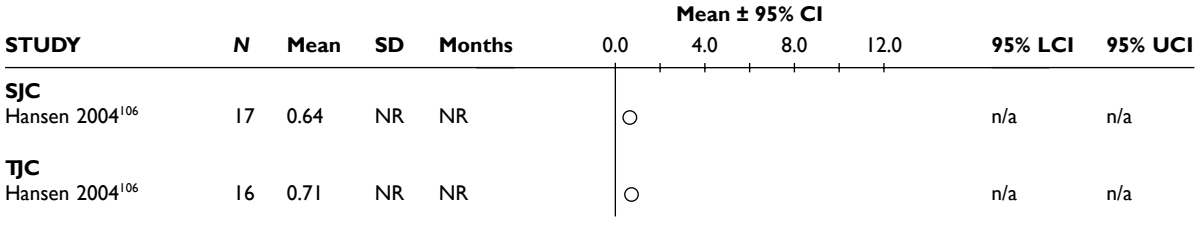


FIGURE 113 Infliximab: percentage change from baseline in SJC and TJC. LCI, lower confidence interval; n/a, not applicable; NR, not reported; SD, standard deviation; UCI, upper confidence interval.

STUDY	N	Mean	SD	Months	95% LCI	95% UCI
Hansen 2004[106]	6	23.80	NR	0	n/a	n/a
Hansen 2004[106]	6	17.10	NR	NR	n/a	n/a

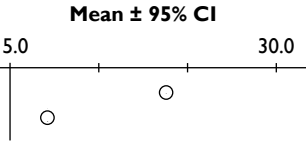


FIGURE 114 Infliximab: mean change from baseline in CRP. LCI, lower confidence interval; n/a, not applicable; NR, not reported; SD, standard deviation; UCI, upper confidence interval.

STUDY	N	Mean	SD	Months	95% LCI	95% UCI
Hansen 2004[106]	11	13.00	NR	0	n/a	n/a
Hansen 2004[106]	11	26.00	NR	NR	n/a	n/a

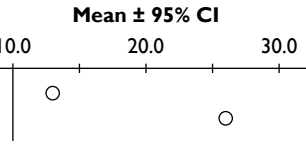


FIGURE 115 Infliximab: mean change from baseline in ESR. LCI, lower confidence interval; n/a, not applicable; NR, not reported; SD, standard deviation; UCI, upper confidence interval.

Infliximab

STUDY	N	Mean	SD	Months	95% LCI	95% UCI
SJC						
Finckh 2007[136]	66	−1.98	4.77	3	−3.15	−0.81
Finckh 2007[136]	66	−3.00	6.22	6	−4.53	−1.47
Finckh 2007[136]	66	−3.60	6.63	9	−5.23	−1.97

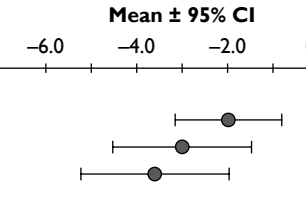


FIGURE 116 TNF inhibitor: SJC. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

STUDY	N	Mean	SD	Months	95% LCI	95% UCI
TJC						
Finckh 2007[136]	66	−2.50	4.14	3	−3.52	−1.48
Finckh 2007[136]	66	−4.00	5.18	6	−5.27	−2.73
Finckh 2007[136]	66	−4.25	6.22	9	−5.78	−2.72

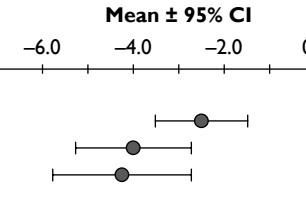


FIGURE 117 TNF inhibitor: TJC. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

STUDY	N	Mean	SD	Months	95% LCI	95% UCI
ESR						
Finckh 2007[136]	66	−3.25	11.40	3	−6.05	−0.45
Finckh 2007[136]	66	−6.00	15.54	6	−9.82	−2.18
Finckh 2007[136]	66	−7.50	16.06	9	−11.45	−3.55

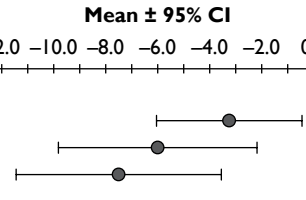


FIGURE 118 TNF inhibitor: mean change from baseline in ESR. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

Tumour necrosis factor inhibitors as a class

Study or subgroup	Rituximab Mean	Rituximab SD	Rituximab Total	Placebo Mean	Placebo SD	Placebo Total	Mean difference IV, Fixed, 95% CI	Mean difference IV, Fixed, 95% CI
REFLEX[124–126]	23.4	29.4	298	−2.5	23.3	201	25.90 (21.26 to 30.54)	

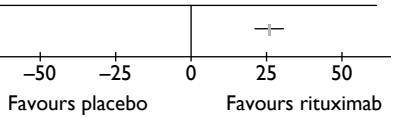


FIGURE 119 Rituximab: patient pain (0–100 mm visual analogue scale) change from baseline at week 24 in the REFLEX RCT.[124–126] SD, standard deviation.

Study or subgroup	Rituximab Mean	Rituximab SD	Rituximab Total	Placebo Mean	Placebo SD	Placebo Total	Mean difference IV, Fixed, 95% CI	Mean difference IV, Fixed, 95% CI
REFLEX[124–126]	−29.5	27.4	298	−6.2	27.1	201	−23.30 (−28.17 to −18.43)	

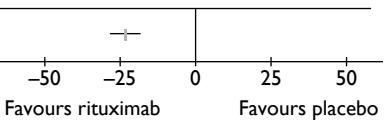


FIGURE 120 Rituximab: physical global function (0–100 mm visual analogue scale) change from baseline at week 24 in the REFLEX RCT.[124–126] SD, standard deviation.

Study or subgroup	Rituximab Mean	Rituximab SD	Rituximab Total	Placebo Mean	Placebo SD	Placebo Total	Mean difference IV, Fixed, 95% CI	Mean difference IV, Fixed, 95% CI
REFLEX[124–126]	−26	30	298	−5.3	22.9	201	−20.70 (−25.35 to −16.05)	

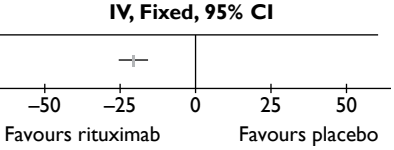


FIGURE 121 Rituximab: patient global function (0–100 mm visual analogue scale) change from baseline at week 24 in the REFLEX RCT.[124–126] SD, standard deviation.

Study or subgroup	Rituximab Mean	Rituximab SD	Rituximab Total	Placebo Mean	Placebo SD	Placebo Total	Mean difference IV, Fixed, 95% CI	Mean difference IV, Fixed, 95% CI
REFLEX[124–126]	−9.14	11.31	298	−0.54	9.84	201	−8.60 (−10.47 to −6.73)	

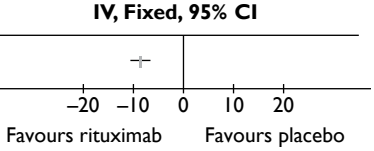


FIGURE 122 Rituximab: change in functional assessment of chronic illness therapy – fatigue (range 0–52) score from baseline at week 24 in the REFLEX RCT.[124–126] SD, standard deviation.

Study or subgroup	Rituximab Mean	Rituximab SD	Rituximab Total	Placebo Mean	Placebo SD	Placebo Total	Mean difference IV, Fixed, 95% CI	Mean difference IV, Fixed, 95% CI
REFLEX[124–126]	−2.1	3.5	298	0	3.6	201	−2.10 (−2.74 to −1.46)	

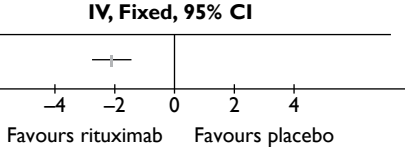


FIGURE 123 Rituximab: mean change in CRP (mg/l) from baseline at week 24 in the REFLEX RCT.[124–126] SD, standard deviation.

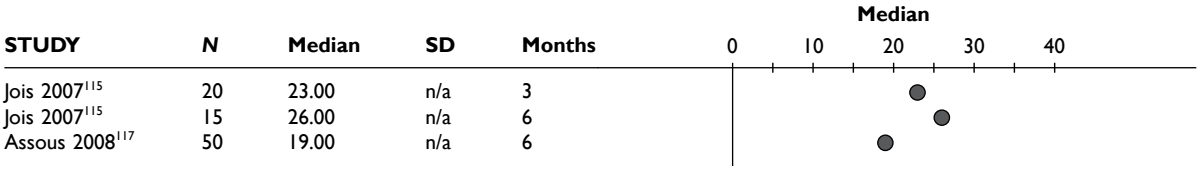


FIGURE 124 Rituximab: median CRP (mg/l) in uncontrolled studies. n/a, not applicable; SD, standard deviation. (Not significant vs baseline for Jois *et al*.[115] study and $p < 0.05$ vs baseline for Assous *et al*.[117])

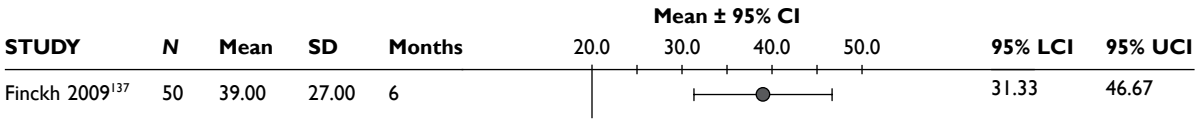


FIGURE 125 Rituximab: mean ESR (mm/h) in uncontrolled studies. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

FIGURE 126 Rituximab: median ESR (mm/h) in uncontrolled studies. ($p < 0.0001$ for both at 3 months and 6 months.) SD, standard deviation.

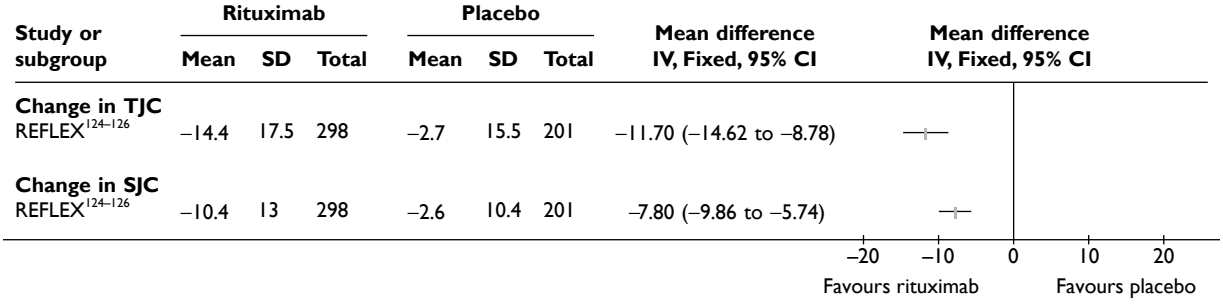


FIGURE 127 Rituximab: change in TJC/SJC from baseline at week 24 in the REFLEX RCT.[124–126] SD, standard deviation.

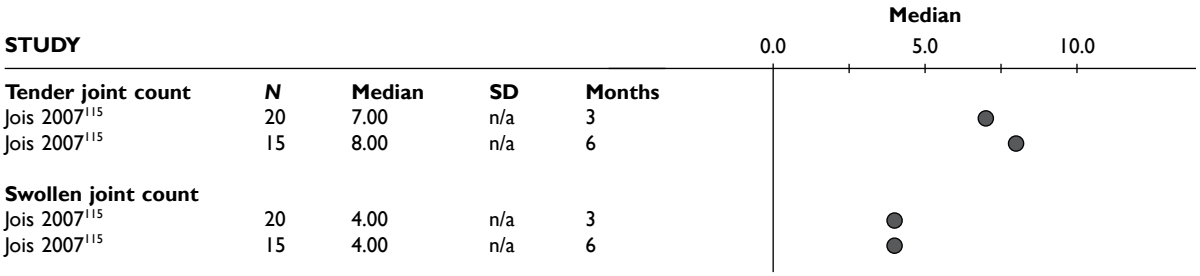


FIGURE 128 Rituximab: median TJC/SJC count in uncontrolled studies ($p < 0.0001$ for 3 months vs baseline and $p < 0.05$ for 6 months vs baseline). SD, standard deviation.

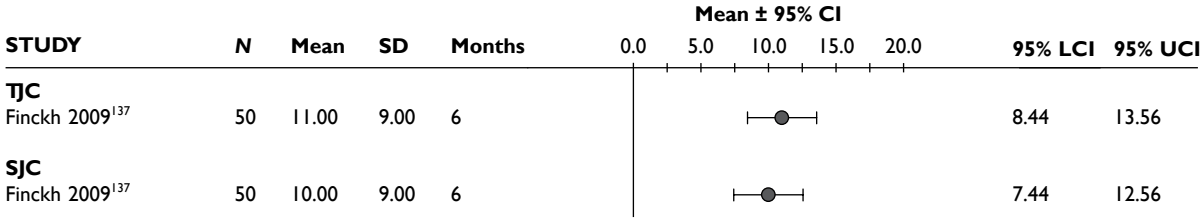


FIGURE 129 Rituximab: mean TJC/SJC count in uncontrolled studies. **Bold type** idicates that the study was an RCT. LCI, lower confidence interval; SD, standard deviation; UCI, upper confidence interval.

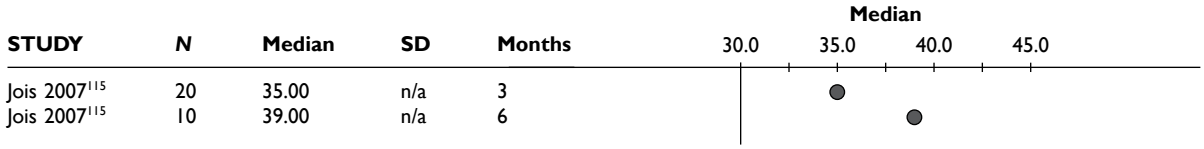


FIGURE 130 Rituximab: median patient global score (0–100 mm visual analogue scale) in uncontrolled studies. SD, standard deviation.

Rituximab

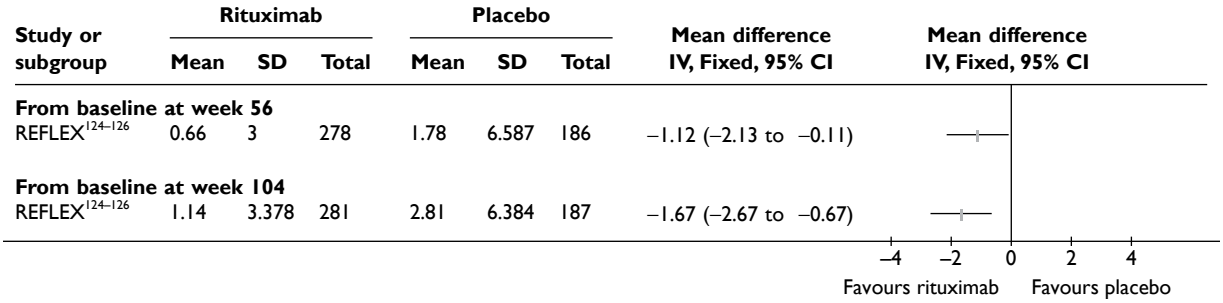


FIGURE 131 Rituximab: Sharp–Genant total score change from baseline in the REFLEX trial[124–126] (data from MS; the SD for that at week 56 was calculated from *p*-value). SD, standard deviation.

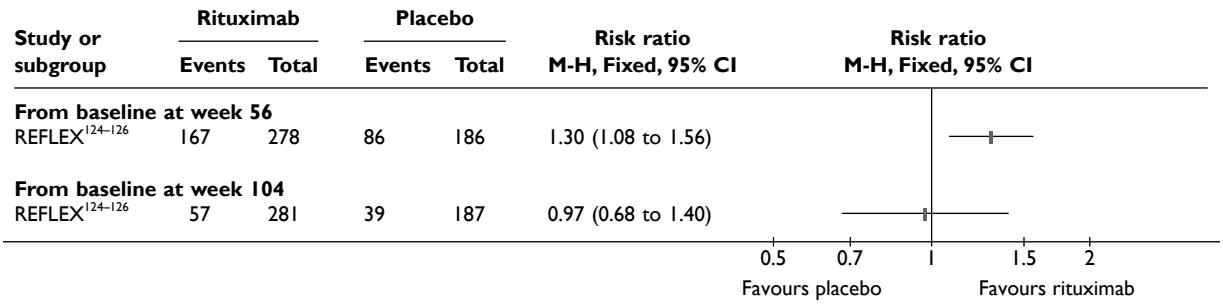


FIGURE 132 Rituximab: percentage of patients with no worsening Sharp–Genant total score from baseline in the REFLEX trial[124–126] (data from MS).

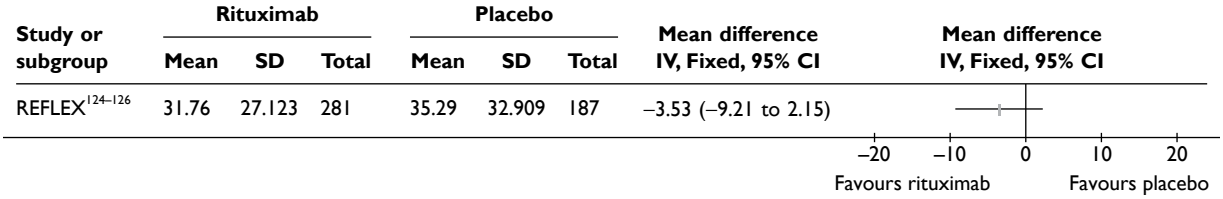


FIGURE 133 Rituximab: Sharp–Genant total score at week 104 in the REFLEX trial[124–126] (data from MS). SD, standard deviation.

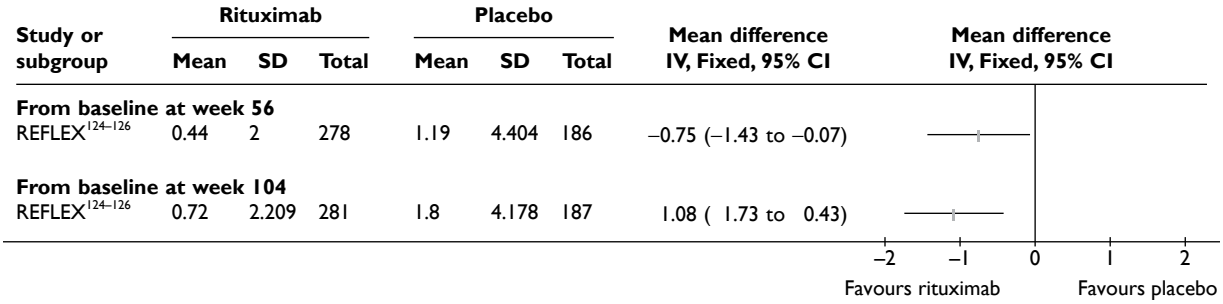


FIGURE 134 Rituximab: erosion score change from baseline in the REFLEX trial[124–126] (data from MS; the SD for that at week 56 was calculated from *p*-value). SD, standard deviation.

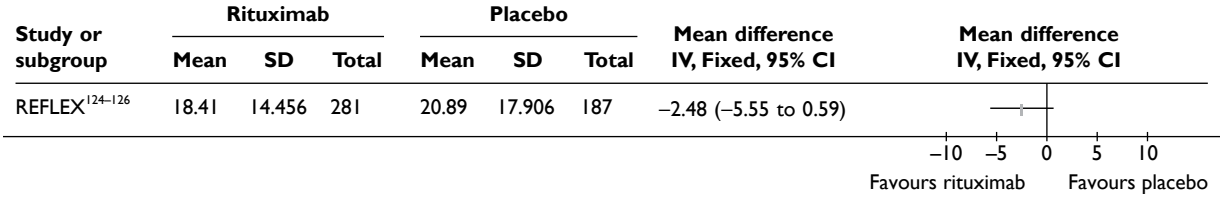


FIGURE 135 Rituximab: erosion scores at week 104 in the REFLEX trial[124–126] (data from MS). SD, standard deviation.

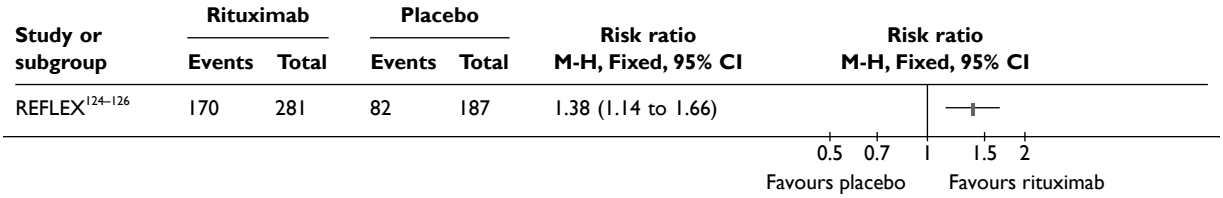


FIGURE 136 Rituximab: percentage of patients with no erosive progression from baseline at week 104 in the REFLEX trial[124–126] (data from MS).

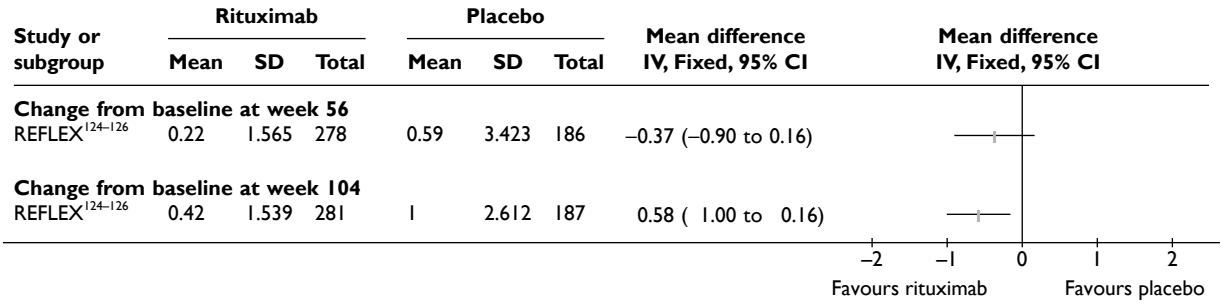


FIGURE 137 Rituximab: joint space narrowing score change from baseline in the REFLEX trial[124–126] (data from MS; for the 56 week the SD was calculated from *p*-value). SD, standard deviation.

Study or subgroup	Rituximab Mean	Rituximab SD	Rituximab Total	Placebo Mean	Placebo SD	Placebo Total	Mean difference IV, Fixed, 95% CI
REFLEX[124–126]	13.35	14.015	281	14.4	16.034	187	−1.05 (−3.87 to 1.77)

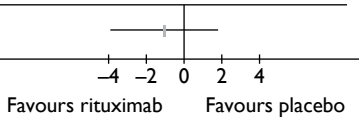
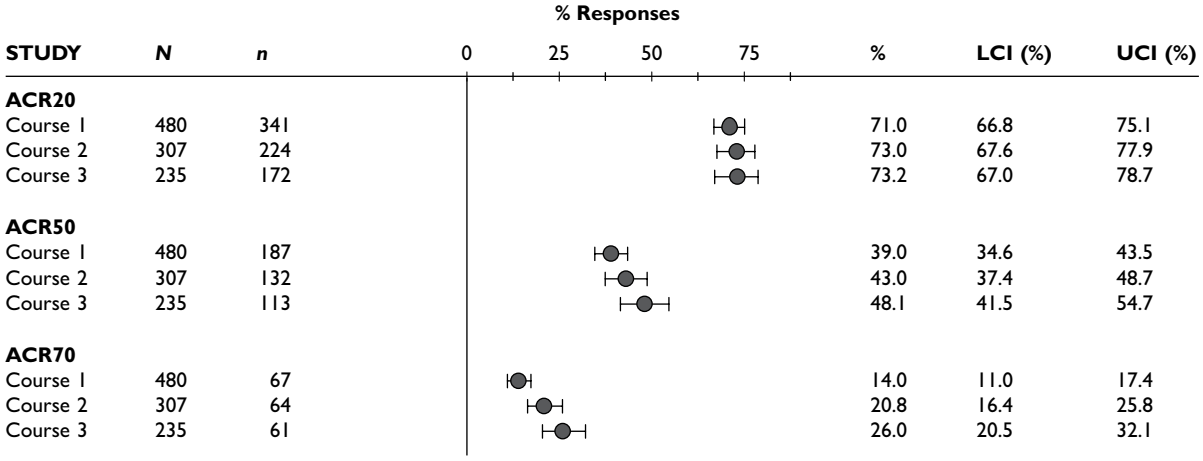


FIGURE 138 Rituximab: joint space narrowing score at week 104 in the REFLEX trial[124–126] (data from MS). SD, standard deviation.

Joint damage data from the manufacturers' submissions

REFLEX extension

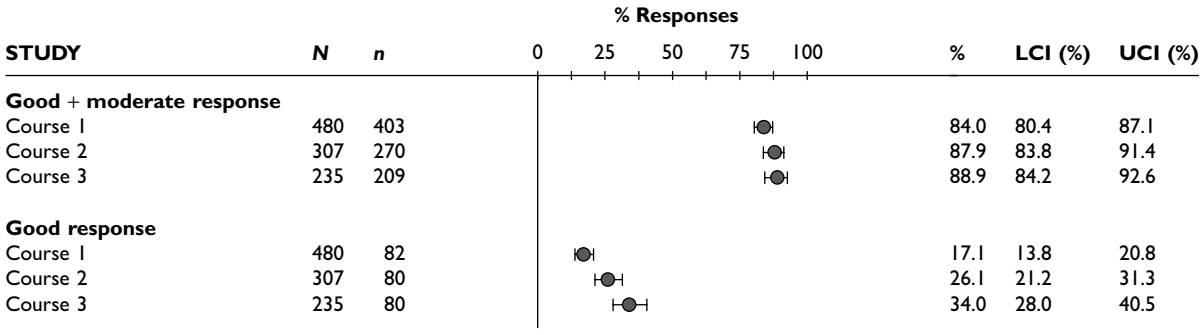
Figure 139 presents ACR response at week 24 after one, two and three RTX treatment courses versus original baseline in the REFLEX trial.[124–126] A similar pattern was seen for each ACR response 24 weeks after each course, with the ACR responses following each course slightly increased with subsequent courses.



STUDY	N	n	%	LCI (%)	UCI (%)
ACR20					
Course 1	480	341	71.0	66.8	75.1
Course 2	307	224	73.0	67.6	77.9
Course 3	235	172	73.2	67.0	78.7
ACR50					
Course 1	480	187	39.0	34.6	43.5
Course 2	307	132	43.0	37.4	48.7
Course 3	235	113	48.1	41.5	54.7
ACR70					
Course 1	480	67	14.0	11.0	17.4
Course 2	307	64	20.8	16.4	25.8
Course 3	235	61	26.0	20.5	32.1

FIGURE 139 Rituximab: ACR response at week 24 after each course versus original baseline (data from MS). LCI, lower confidence interval; UCI, upper confidence interval.

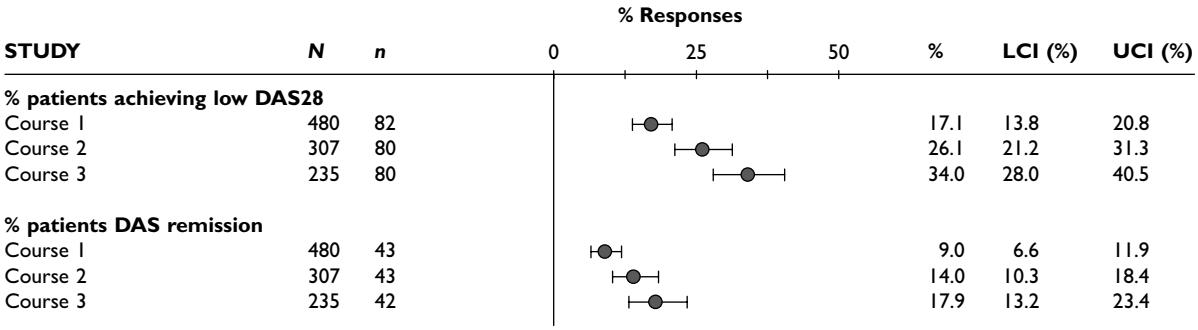
Figure 140 presents EULAR response at week 24 after one, two and three courses of RTX versus original baseline of the REFLEX trial.[124–126] The percentage of patients achieving moderate plus good response and good response alone increased with each treatment course (from 84% to 87.9% to 88.9% and from 17.1% to 26.1% to 28%, respectively).



STUDY	N	n	%	LCI (%)	UCI (%)
Good + moderate response					
Course 1	480	403	84.0	80.4	87.1
Course 2	307	270	87.9	83.8	91.4
Course 3	235	209	88.9	84.2	92.6
Good response					
Course 1	480	82	17.1	13.8	20.8
Course 2	307	80	26.1	21.2	31.3
Course 3	235	80	34.0	28.0	40.5

FIGURE 140 Rituximab: EULAR responses 24 weeks after each course versus original baseline (data from MS). LCI, lower confidence interval; UCI, upper confidence interval.

Figure 141 presents the percentage of patients achieving DAS28 low disease activity or remission at week 24 after course one, two and three versus original baseline of the REFLEX trial.[124–126] Improvement for both was observed following subsequent courses (from 17.1% to 26.1% to 34% and from 9% to 14% to 13.2%, respectively).

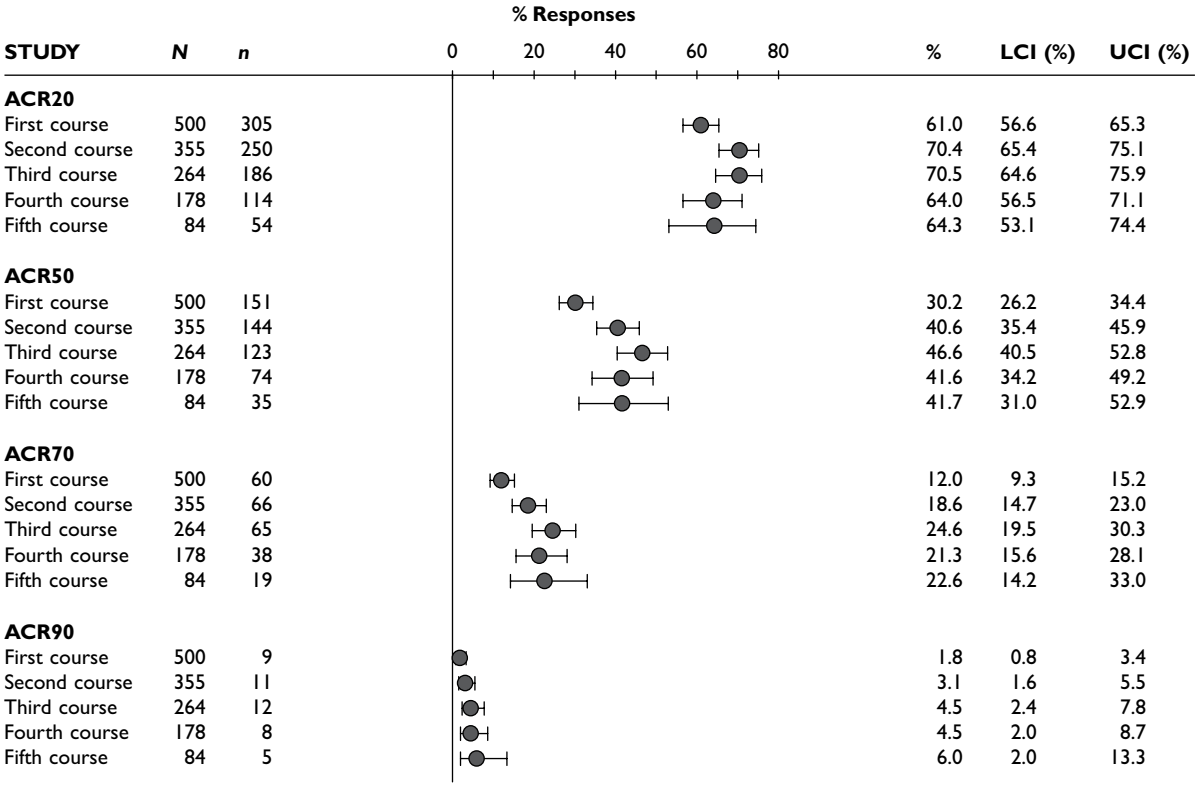


STUDY	N	n	%	LCI (%)	UCI (%)
% patients achieving low DAS28					
Course 1	480	82	17.1	13.8	20.8
Course 2	307	80	26.1	21.2	31.3
Course 3	235	80	34.0	28.0	40.5
% patients DAS remission					
Course 1	480	43	9.0	6.6	11.9
Course 2	307	43	14.0	10.3	18.4
Course 3	235	42	17.9	13.2	23.4

FIGURE 141 Rituximab: percentage of patients achieving DAS28 low disease activity at week 24 after each course versus original baseline (data from MS). LCI, lower confidence interval; UCI, upper confidence interval.

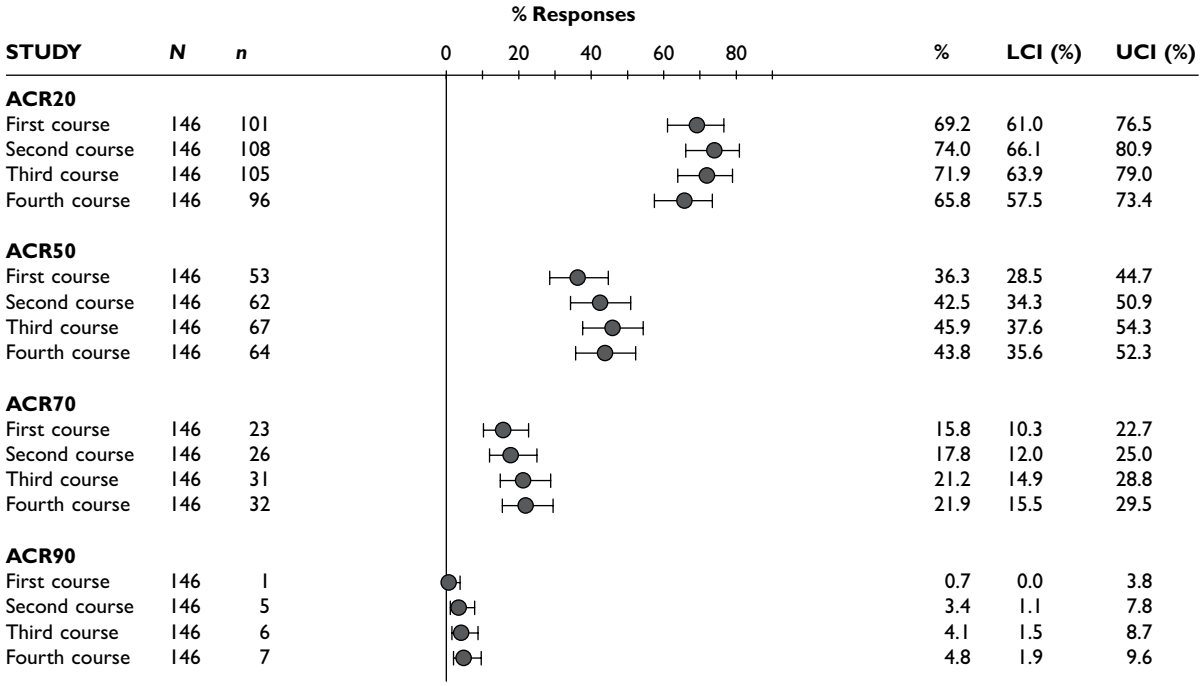
Pooled analysis data (manufacturer's submission)

Figure 142 presents ACR responses for four or five courses and *Figure 143* presents ACR responses for three or four courses of RTX 24 weeks after each course. The overall pattern was that there was an improvement from the first to the second course and then maintained through the subsequent courses. Observed data on EULAR responses for four or five courses at 24 weeks after each course showed a similar pattern to ACR responses (*Figure 144*).



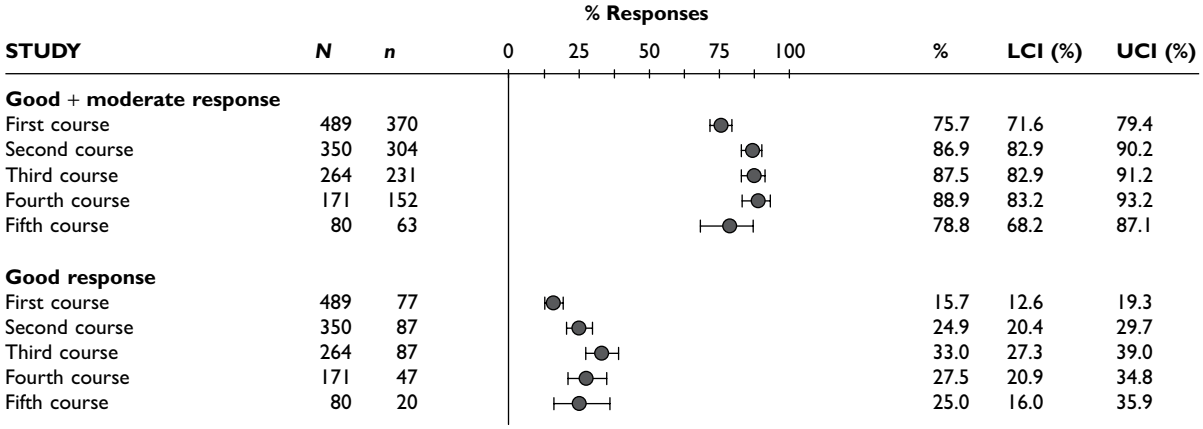
STUDY	N	n	%	LCI (%)	UCI (%)
ACR20					
First course	500	305	61.0	56.6	65.3
Second course	355	250	70.4	65.4	75.1
Third course	264	186	70.5	64.6	75.9
Fourth course	178	114	64.0	56.5	71.1
Fifth course	84	54	64.3	53.1	74.4
ACR50					
First course	500	151	30.2	26.2	34.4
Second course	355	144	40.6	35.4	45.9
Third course	264	123	46.6	40.5	52.8
Fourth course	178	74	41.6	34.2	49.2
Fifth course	84	35	41.7	31.0	52.9
ACR70					
First course	500	60	12.0	9.3	15.2
Second course	355	66	18.6	14.7	23.0
Third course	264	65	24.6	19.5	30.3
Fourth course	178	38	21.3	15.6	28.1
Fifth course	84	19	22.6	14.2	33.0
ACR90					
First course	500	9	1.8	0.8	3.4
Second course	355	11	3.1	1.6	5.5
Third course	264	12	4.5	2.4	7.8
Fourth course	178	8	4.5	2.0	8.7
Fifth course	84	5	6.0	2.0	13.3

FIGURE 142 Rituximab: ACR responses for five courses of treatment 24 weeks after each course (all patients; observed data; data from MS). LCI, lower confidence interval; UCI, upper confidence interval.



STUDY	N	n	%	LCI (%)	UCI (%)
ACR20					
First course	146	101	69.2	61.0	76.5
Second course	146	108	74.0	66.1	80.9
Third course	146	105	71.9	63.9	79.0
Fourth course	146	96	65.8	57.5	73.4
ACR50					
First course	146	53	36.3	28.5	44.7
Second course	146	62	42.5	34.3	50.9
Third course	146	67	45.9	37.6	54.3
Fourth course	146	64	43.8	35.6	52.3
ACR70					
First course	146	23	15.8	10.3	22.7
Second course	146	26	17.8	12.0	25.0
Third course	146	31	21.2	14.9	28.8
Fourth course	146	32	21.9	15.5	29.5
ACR90					
First course	146	1	0.7	0.0	3.8
Second course	146	5	3.4	1.1	7.8
Third course	146	6	4.1	1.5	8.7
Fourth course	146	7	4.8	1.9	9.6

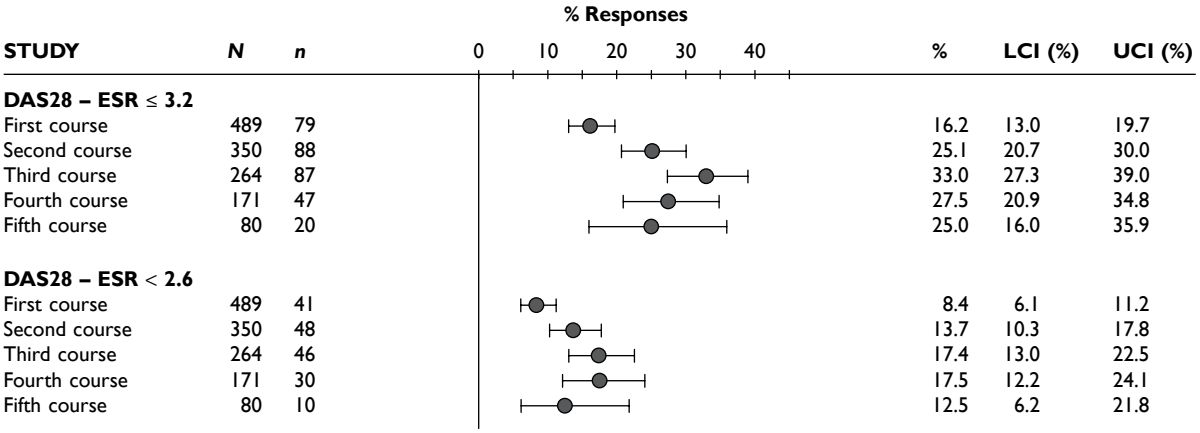
FIGURE 143 Rituximab: ACR responses for three or four course (24 weeks) after each course (within-patients, within-visit comparisons, observed data, $n = 146$; data from MS). LCI, lower confidence interval; UCI, upper confidence interval.



STUDY	N	n	%	LCI (%)	UCI (%)
Good + moderate response					
First course	489	370	75.7	71.6	79.4
Second course	350	304	86.9	82.9	90.2
Third course	264	231	87.5	82.9	91.2
Fourth course	171	152	88.9	83.2	93.2
Fifth course	80	63	78.8	68.2	87.1
Good response					
First course	489	77	15.7	12.6	19.3
Second course	350	87	24.9	20.4	29.7
Third course	264	87	33.0	27.3	39.0
Fourth course	171	47	27.5	20.9	34.8
Fifth course	80	20	25.0	16.0	35.9

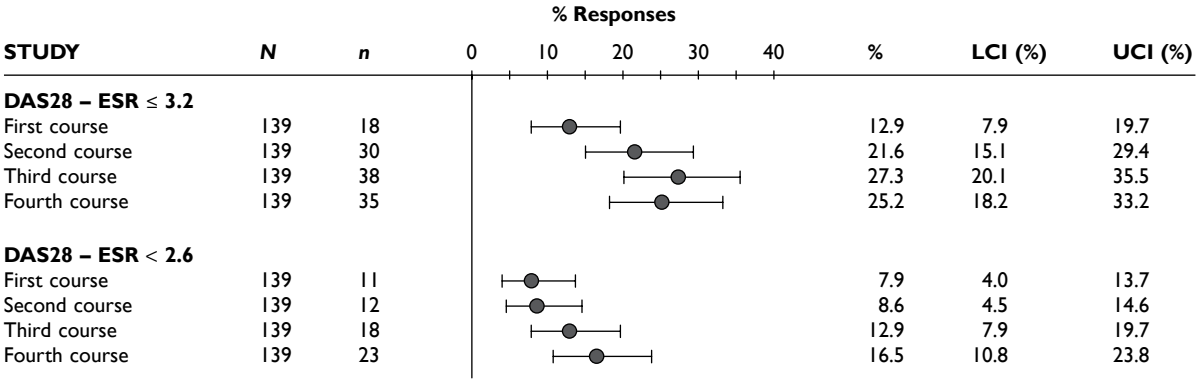
FIGURE 144 Rituximab: EULAR response rates for four or five courses (week 24 after each course, all patients, observed data; data from MS). LCI, lower confidence interval; UCI, upper confidence interval.

The patterns for the percentage of patients with low disease activity (defined as DAS28-ESR less than or equal to 3.2) and with remission (defined as DAS28-ESR less than 2.6) for four or five courses at week 24 after each course, and for data on three or four courses at week 24 after each course, were similar, with a improvement from first to second course and to third course and then generally maintained with subsequent courses (*Figures 145* and *146*).



STUDY	N	n	%	LCI (%)	UCI (%)
DAS28 – ESR ≤ 3.2					
First course	489	79	16.2	13.0	19.7
Second course	350	88	25.1	20.7	30.0
Third course	264	87	33.0	27.3	39.0
Fourth course	171	47	27.5	20.9	34.8
Fifth course	80	20	25.0	16.0	35.9
DAS28 – ESR < 2.6					
First course	489	41	8.4	6.1	11.2
Second course	350	48	13.7	10.3	17.8
Third course	264	46	17.4	13.0	22.5
Fourth course	171	30	17.5	12.2	24.1
Fifth course	80	10	12.5	6.2	21.8

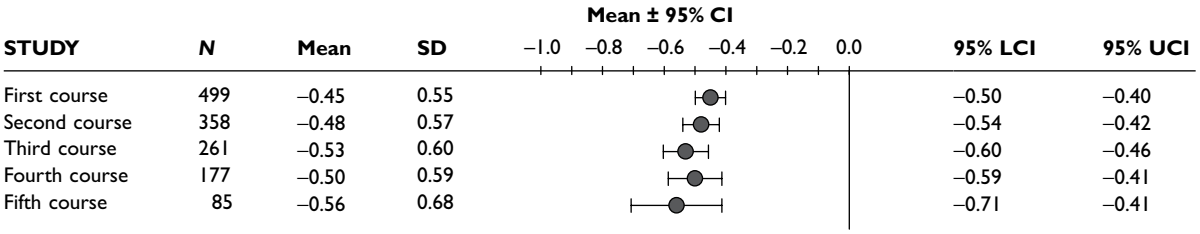
FIGURE 145 Rituximab: percentage of patients with low disease activity (DAS28-ESR ≤3.2) and with remission of disease activity (DAS28-ESR <2.6) for four or five courses (week 24 after each course, all patients, observed data; data from MS). LCI, lower confidence interval; UCI, upper confidence interval.



STUDY	N	n	%	LCI (%)	UCI (%)
DAS28 – ESR ≤ 3.2					
First course	139	18	12.9	7.9	19.7
Second course	139	30	21.6	15.1	29.4
Third course	139	38	27.3	20.1	35.5
Fourth course	139	35	25.2	18.2	33.2
DAS28 – ESR < 2.6					
First course	139	11	7.9	4.0	13.7
Second course	139	12	8.6	4.5	14.6
Third course	139	18	12.9	7.9	19.7
Fourth course	139	23	16.5	10.8	23.8

FIGURE 146 Rituximab: percentage of patients with low disease activity (DAS28-ESR ≤3.2) and with remission of disease activity (DAS28-ESR <2.6) for three or four course (week 24 after each course, all patients, observed data; data from MS). LCI, lower confidence interval; UCI, upper confidence interval.

Figure 147 presents the change from original baseline of the REFLEX trial[124–126] in HAQ for four or five courses 24 weeks after each course and *Figure 148* presents the percentage of patients achieving minimally important clinical difference, i.e. a decrease in HAQ score of greater than or equal to 0.22 from baseline, for four or five courses 24 weeks after each course. Both the change in HAQ score and the percentage of patients achieving a clinically meaningful decrease in HAQ score were maintained over treatment courses of RTX.



STUDY	N	Mean	SD	95% LCI	95% UCI
First course	499	−0.45	0.55	−0.50	−0.40
Second course	358	−0.48	0.57	−0.54	−0.42
Third course	261	−0.53	0.60	−0.60	−0.46
Fourth course	177	−0.50	0.59	−0.59	−0.41
Fifth course	85	−0.56	0.68	−0.71	−0.41

FIGURE 147 Rituximab: change from original baseline in HAQ end points for four or five courses (week 24 after each course, all patients, observed data; data from MS). LCI, lower confidence interval; UCI, upper confidence interval.

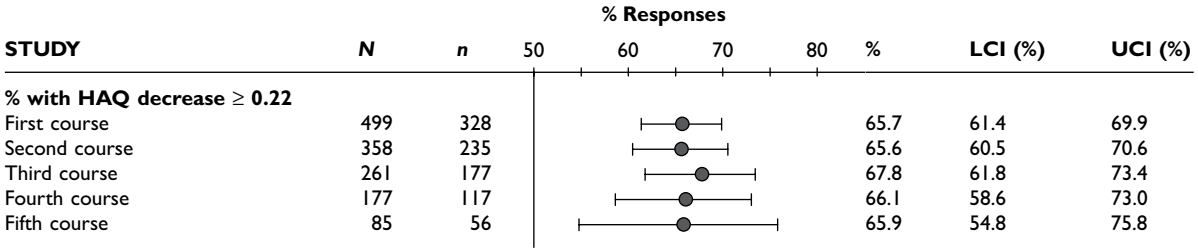


FIGURE 148 Rituximab: percentage of patients with HAQ decrease ≥0.22 from original baseline for four or five courses (week 24 after each course, all patients, observed data; data from MS). LCI, lower confidence interval; UCI, upper confidence interval.

Abatacept

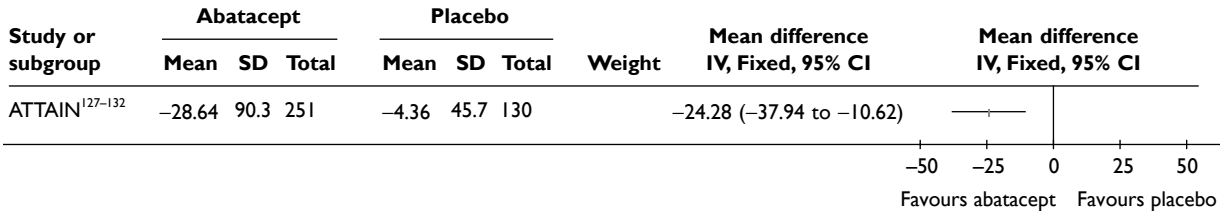


FIGURE 149 Abatacept: change in pain score (visual analogue scale) in the ATTAIN RCT[127–132] at 6 months. SD, standard deviation.

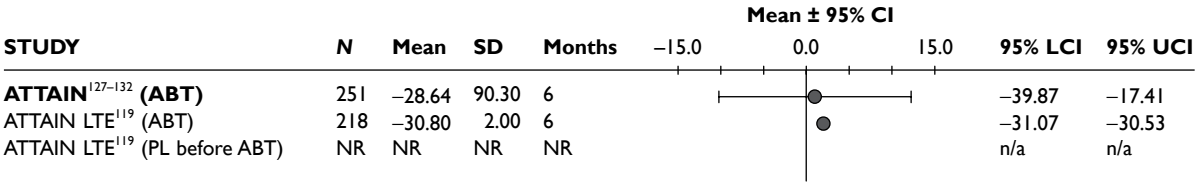


FIGURE 150 Abatacept: change in pain score (visual analogue scale) in uncontrolled studies. LCI, lower confidence interval; n/a, not applicable; NR, not reported; SD, standard deviation; UCI, upper confidence interval.

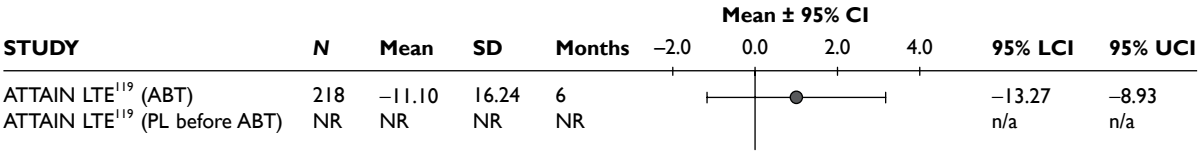


FIGURE 151 Abatacept: change in sleep score in uncontrolled studies at 6 months. LCI, lower confidence interval; n/a, not applicable; NR, not reported; SD, standard deviation; UCI, upper confidence interval.

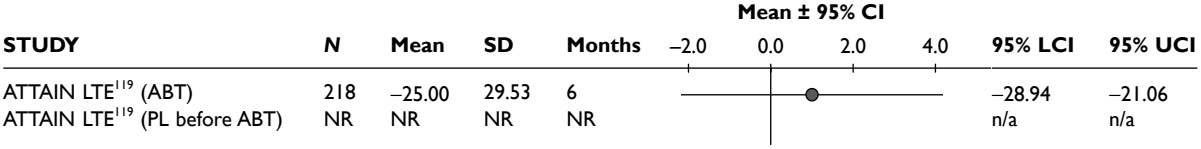


FIGURE 152 Abatacept: change in fatigue score in uncontrolled studies at 6 months. LCI, lower confidence interval; n/a, not applicable; NR, not reported; SD, standard deviation; UCI, upper confidence interval.

Appendix 11

Survey of West Midlands rheumatologists

A survey of rheumatologists in the West Midlands was conducted in June 2009 and July 2009 to investigate current practice and clinicians' preferences for treatment options in rheumatoid arthritis.

Methods

In the beginning of June a questionnaire was sent to a convenience sample of 55 rheumatologists by email (*Figure 153*).

1. Which DMARD(s), in addition to MXT, do you normally try before using a TNF inhibitor?

2. Which is your preferred first choice TNF inhibitor, if any?

3. In people not responding adequately to a TNF inhibitor, assuming that another TNF inhibitor, rituximab, abatacept, and tocilizumab were all available and not restricted by NICE, which DMARDs (including non-biologic agents) would you next try (jot down an ideal sequence of individual or combinations you prefer, in sequence, and ignore issues of local logistics and of patients co-morbidity)

a) first .....

b) second (if the first fails) .....

c) third (etc.) .....

d) fourth .....

e) fifth (and beyond, continue as long as your imagination or patience allows)

.....

4. Please write here any general comments or thoughts

FIGURE 153 Survey of West Midlands rheumatologists.

Responses were collected until early July, when a reminder together with the results of the survey so far was sent. Responses received afterwards were included in the results.

Owing to the overall variability it was not possible to determine in any way if the three responses received after the reminder were influenced by the knowledge of the early results.

Results

Twenty-four rheumatologists replied before the reminder email. Three additional responses were received after the reminder was sent out. The overall response rate was 49%.

For drugs used in addition to MTX before the initiation of the first TNF inhibitor responses often included combinations of multiple conventional DMARDs or different therapeutic options. Sulfasalazine alone or in combination with other DMARDs was the most frequently mentioned

DMARD (in 22 responses) used before the initiation of the first TNF inhibitor. Leflunomide was mentioned in 17 responses and hydroxychloroquine in 10. Five respondents mentioned the use of steroids.

Results for the first TNF inhibitor and following treatment options are presented in *Figure 154*. The highest number of respondents (nine) left the choice of the first TNF inhibitor to the patient. Seven would chose ADA and one indicated that this drug was most often chosen by patients. Etanercept was the preferred first TNF inhibitor for six respondents; however, three would ultimately leave the choice to their patient. The remaining four would choose either ADA or ETN (two because of involvement in a clinical trial).

After the failure of the first TNF inhibitor, 17 respondents would try a second one (only six were specific and their preferences were – ADA in four and ETN in two cases). Nine respondents would try RTX as a second-line biologic agent and one TOC.

There was more variability in the following lines of treatment and preferences depended on what has been tried before. After the failure of a second TNF inhibitor, ten respondents would try RTX, five TOC, one ADA and one LEF. After the failure of RTX (following first TNF inhibitor) six respondents would try a second TNF inhibitor, two would try TOC and one ABT. One respondent who would try TOC after the failure of the first TNF inhibitor would choose RTX as the next therapeutic option.

For the next line of treatment see *Figure 154*. Results for the subsequent treatment options are not reported because of their high variability.

The comments from respondents included a number of issues referring both to current practice and to proposed research:

- Different factors might influence choice of drug, such as:
  - previous or possible tuberculosis
  - risk of infection
  - comorbidities
  - primary versus secondary failure
  - seropositive versus negative patients
  - intolerance versus inefficacy
  - ethnicity (ETN preferred in Asian patients)
  - 'needle-phobia'.
- Practice is frequently tailored to the individual patient (pattern of disease, side-effect risks, etc.).
- Going back to a TNF inhibitor already used could be considered.
- For some patients receiving biologic treatments, adjunct DMARDs other than MTX could be considered.
- Switching TNF inhibitors before the 3-month NICE deadline could be considered if the patient showed little response.
- A combination of TNF inhibitors could be considered.

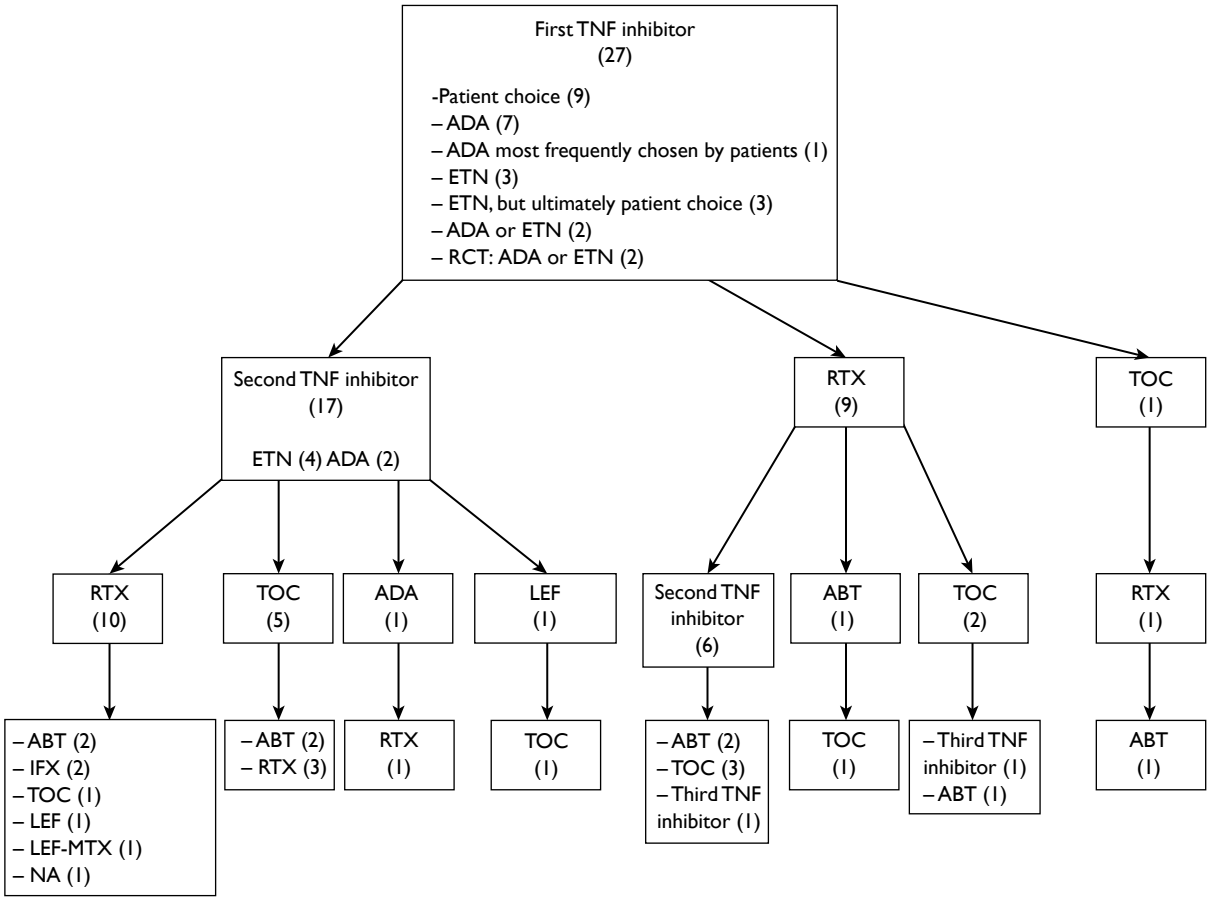


FIGURE 154 Survey of West Midlands rheumatologists: results. Numbers in brackets are the number of respondents selecting and option.

Appendix 12

Withdrawals from treatment with tumour necrosis factor inhibitors

Withdrawal from treatment with second-line tumour necrosis factor inhibitor (British Society for Rheumatology Biologics Registry data)

Updated BSRBR model data[163] provided Kaplan–Meier (K–M) plots for survival in treatment for four groups of patients receiving second-line TNF inhibitors as follows: (i) withdrew from first-line TNF inhibitor for lack of efficacy and from second-line TNF inhibitor for lack of efficacy; (ii) withdrew from first-line TNF inhibitor for lack of efficacy and from second line TNF inhibitor for AEs; (iii) withdrew from first-line TNF inhibitor for AEs and from second-line TNF inhibitor for lack of efficacy; (iv) withdrew from first-line TNF inhibitor for AEs and from second-line TNF inhibitor for AEs.

The proportion lost to treatment at 3-month time points in each category was read from the graphs in the BSRBR submission and the absolute number lost calculated using $N = 995$ for first-line withdrawal through lack of efficacy and $N = 1{,}882$ for first-line withdrawal due to AEs. The proportion of patients withdrawing for any reason was then estimated and the proportion remaining in treatment plotted (data points in *Figure 155*). A Weibull distribution (time in years) was fitted to the data [scale parameter (lambda) 0.441555; standard error (SE) 0.00958300], shape parameter (gamma) 0.7008 (SE 0.033681) labelled BSRBR Weibull fit in *Figure 155* (extrapolation to 25 years is shown in the inset).

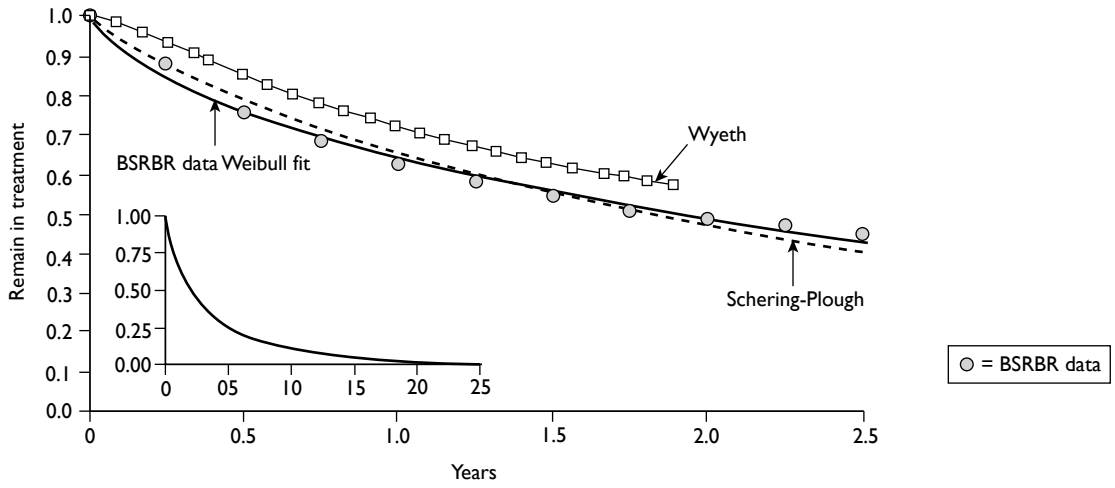


FIGURE 155 Continuation in second-line TNF inhibitor.

Comparison with manufacturers' submissions

The Schering-Plough Ltd (infliximab) submission[165] provided Weibull parameters for treatment withdrawal that were also based on BSRBR data; the parameters are shown below.

Log(scale) 3.529 (time in months)

Log(shape) –0.19 (time in months)

Assuming log(scale) in the table above refers to 'log β' where $\beta = (1/\lambda) \wedge [1/\gamma]$, and survival $= \exp[-(t \times \beta) \wedge \gamma]$, then lambda = 0.054 and gamma = 0.827 and the fitted curve labelled Schering-Plough Ltd in *Figure 155* is generated (and can be seen to be very similar to the review group's fit).

The Wyeth Pharmaceuticals submission[226] modelled withdrawal from treatment using a 'shared frailty' model and this is also represented in *Figure 155*.

Withdrawal from second-line treatment according to tumour necrosis factor inhibitor

According to analysis of Danish Registry for Biologic Therapies in Rheumatology (DANBIO) data withdrawal from first line TNF inhibitor occurs at rates that are statistically significantly different between the three TNF inhibitors, *Table 112* provides the reported HRs and 95% CIs (Hetland *et al.*[227]).

TABLE 112 Hazard ratios for withdrawal from first-line TNF inhibitors (DANBIO data)

Comparison	HR	HR 95% CIs	Weibull fit HR
ADA vs ETN	1.35	1.13 to 1.61	1.28
IFX vs ETN	2.10	1.70 to 2.59	1.80
IFX vs ADA	1.56	1.26 to 1.94	1.41

ADA, adalimumab; ETN, etanercept; IFX, infliximab.

It may be reasonable to expect that similar differences might apply for second line TNF inhibitors.

Data were extracted from the K–M graph for each TNF inhibitor published for the Danish registry.[227] These were fitted with Weibull distributions (*Figure 156*) and survivors then combined for each drug (according to number of patients given each TNF inhibitor) so as to provide overall survival ($N = 2{,}935$), and this in turn was fitted with a Weibull distribution.

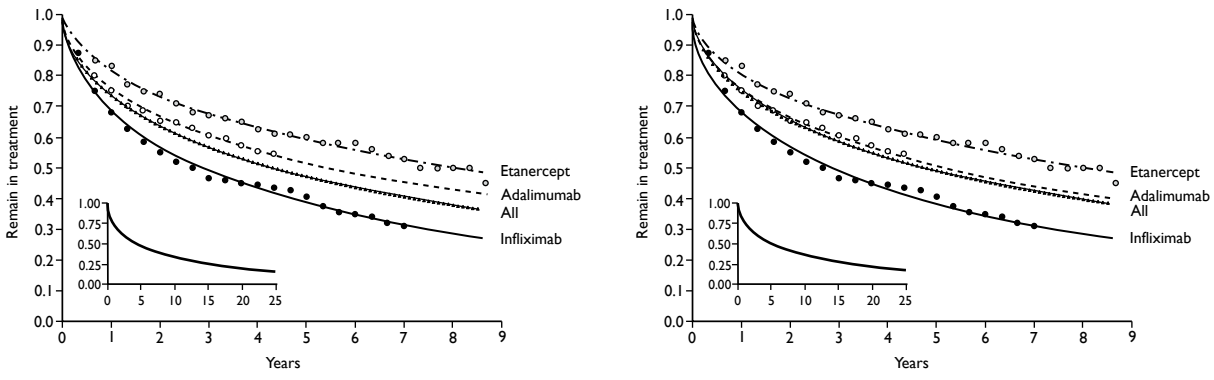


FIGURE 156 Withdrawal from first-line TNF inhibitors (DANBIO data with Weibull fits).

The shape parameters for the Weibull fits were similar and therefore it was considered reasonable to average these and apply the same shape parameter for each drug and for overall survival. Because the BSRBR first-line withdrawal data were derived using equal numbers of patients (~4,000) treated with each TNF inhibitor the shape parameters for the DANBIO data were combined to give an unweighted average. Using this 'common' shape parameter (0.5595) the data were again fitted with Weibull distributions, providing the fits shown in *Figure 156*; the overall survival then assumed that equal numbers received each of the three TNF inhibitors; this allows a comparison of DANBIO and BSRBR first-line withdrawal data (see below).

The HRs (ratio of scale parameters) for comparison of TNF inhibitors using these Weibull fits were within the HR 95% CIs reported for the Danish registry data (note: contact with the lead author confirmed that the published HRs were reversed for ADA versus ETN and IFX versus ADA; this has been corrected in *Table 112*. Relative to all patients (equal mixture) the HRs for each TNF inhibitor were calculated as follows: ETN versus all, 0.751; ADA versus all, 0.958; IFX versus all, 1.353.

When these HRs are applied to the Weibull fit of BSRBR data[163] for continuation of second-line treatment, the drug-specific rates of withdrawal over 25 years are as shown in *Figure 157*.

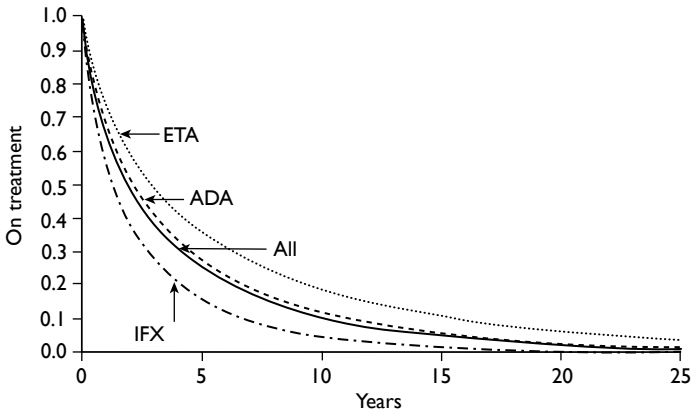


FIGURE 157 Estimated continuation of second-line treatment according to TNF inhibitor.

The Danish National Registry and the British Society for Rheumatology Biologics Registry withdrawal rates from first-line tumour necrosis factor inhibitor

Data for first-line withdrawal were extracted from the UK BSRBR submission[163] and fitted with Weibull distributions in which the shape parameter was or was not fixed to that for overall survival derived from the DANBIO data (0.5595, see above). Extrapolations to 25 years were compared between UK and Danish first-line treatments and between first-line and second-line treatments (*Figure 158*).

Additional sources of evidence

Several additional sources were identified with potentially relevant information on withdrawal from the different TNF inhibitors; these are listed in *Table 113*.

Except for the DANBIO registry data[227] the studies do not provide the information required (K–M plots) to easily compare withdrawal rates between different TNF inhibitors, the main reasons being mixed analysis of first- and second-line withdrawal, mixed populations [rheumatoid arthritis (RA) only a subpopulation, or outcome measure a combination of switching and of dose escalation.[218] The German study[229] does provide information for ETN

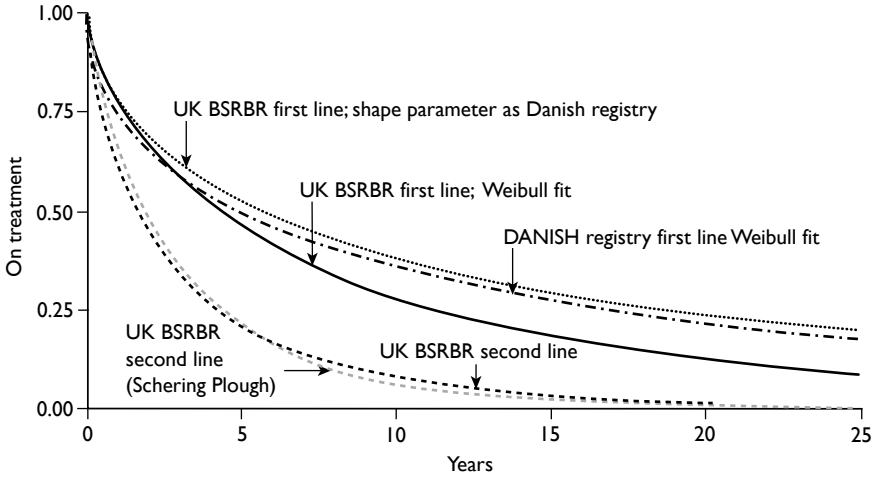


FIGURE 158 Modelled survival in treatment with first- and second-line TNF inhibitors.

and ADA but follow-up was insufficient to see any difference developing. Wolfe and Michaud[230] reported median survival on second-line TNF inhibitor. These results (*Table 114*) compare reasonably well with the median survival for each TNF inhibitor calculated as described above and shown in *Figure 157*.

In general, the data from these studies are consistent with the DANBIO study in that continuation with ETN appears to be superior to that with infliximab and continuation with ADA treatment being intermediate.

TABLE 113 Studies reporting withdrawal rates from TNF inhibitors

Study country	Population (*n*) TNF inhibitors	First-line/second-line withdrawal	Findings	Comment
DANBIO; Hetland 2009;[227] Denmark	RA [National registry] (2,935) IFX, ETN and ADA	Withdrawal from first line	Withdrawal more likely for IFX than ADA and for ADA than ETN	Separate data for withdrawal from first-line treatment with each TNF inhibitor
Finckh 2006;[231] Switzerland	RA only (1,198) IFX, ETN and ADA	Mixed, not differentiated	No difference between IFX, ETN and ADA after adjustment for RF positivity, baseline DAS28, HAQ, failure of previous TNF inhibitor	Not useful for first-line or second-line withdrawal for RA
Duclos 2006;[232] France	Mix of RA [57%] and SpA [one centre] (770) IFX, ETN and ADA	Mixed, not differentiated	No difference between TNF inhibitors. Retention longer for first line vs second line (HR 2.17; 95% CI 1.82 to 2.58, $p<0.0001$) and better if concomitant DMARD	Not useful for first-line or second-line withdrawal for RA
Gomez-Reino 2006;[108] Spain	Mixed [68% RA] (4,706) IFX, ETN and ADA	Both first-line and second-line differentiated	Retention longer for first line vs second line, and for second line vs third line. Second-line retention better if first-line failure was for AEs rather than lack of efficacy. Retention *n* IFX influenced by availability of ETN. Second-line retention better after switch to ETN from IFX than if to switch to IFX from ETN	Not useful for first-line or second-line withdrawal for RA
Vollenhoven 2005;[233] Sweden	'Rheumatic diseases' (128) IFX, ETN and ADA	Second-line withdrawal for lack of efficacy	Less withdrawal from ETN than from IFX; ADA data immature	Not useful for first-line or second-line withdrawal for RA

TABLE 113 Studies reporting withdrawal rates from TNF inhibitors *(continued)*

Study country	Population (*n*) TNF inhibitors	First-line/second-line withdrawal	Findings	Comment
Kristensen 2006;[234] Sweden	RA only (1,161) IFX and ETN	First line; separate analyses according to ± concomitant DMARD and ± MTX	Retention better with ETN than IFX Better retention if patient also receives MTX	K–M data for three subgroups; overall withdrawal from first line with each TNF inhibitor difficult to compute
Zink 2005;[228] Germany	RA (854) IFX and ETN	First line	No statistically significant difference in retention at 12 months: 65.4% for IFX and 68.6% for ETN	Data too immature to draw conclusions
Curtis 2009;[228] USA	RA (11,903) IFX, ETN and ADA	Withdrawal from first line or dose escalation	Hazard ratio for switch from TNF inhibitor (to other DMARD) or dose escalation: IFX vs ETN 6.29 (5.82 to 6.81) ADA vs ETN 1.18 (1.08 to 1.30)	Combines discontinuation and dose escalation
Wolfe and Michaud 2007;[230] USA	RA (4,915) IFX, ETN and ADA	Mixed and second line	Median continuation (years): For first and second line: ADA 3.0, ETN 5.5, IFX 4.5. For second line: ADA 2.0, ETN 2.5, IFX 2.5	K–M plots not supplied

SpA, spondyloarthropathies.

TABLE 114 Median survival for second-line TNF inhibitors

	Median survival second line (years)	
TNF inhibitor	Reported by Wolfe and Michaud 2007[230]	Estimated (as *Figure 157*)
ADA	2	2.02
ETN	2.5	2.86
IFX	2.5	1.24
All[a]	2.36	1.90

a Weighted average according to number of patients receiving each TNF inhibitor.

Appendix 13

Scatter plots for comparisons among biologics in the reference case

This appendix (*Figure 159*) contains the cost-effectiveness scatter plots for the 10 comparisons between biologic treatments in the reference case. The comparisons between biologics and conventional DMARDs are shown in *Chapter 4*, *Reference case*.

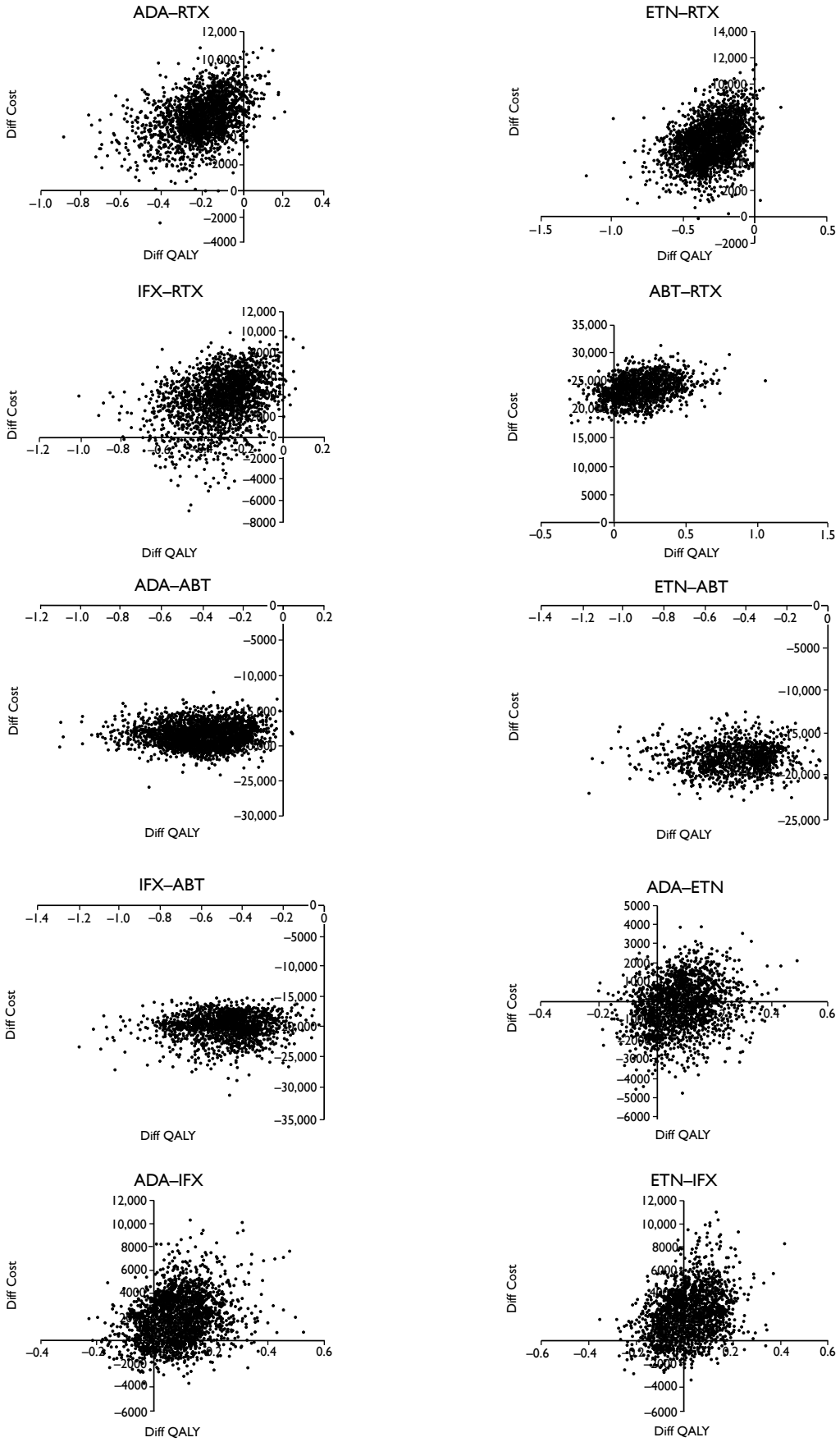


FIGURE 159 Cost-effectiveness scatter plots for comparisons between biologic treatments in the reference case. Diff, difference.

Appendix 14

Scenario analyses

The following scenarios were considered in addition to the reference case analysis. The section headings correspond to the abbreviated descriptions used in *Chapter 4*, *Scenario analysis*. In each case, any parameters not mentioned in the description of the scenario remain as in the reference case analysis.

Vary time on tumour necrosis factors inhibitors

In this case, the time to withdrawing TNF inhibitors treatments was changed to give the same relative risk as for their use as first biologic agents. The *b* parameters from *Table 80* (for reference case) were changed as follows:

Treatment	Reference case *b* parameter (point estimate)	New *b* parameter (point estimate)
ADA	3.211	3.413
ETN	3.211	4.831
IFX	3.211	2.086

Results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	75,900	69,800	82,200	2.92	–2.07	7.92
ETN	82,700	76,000	89,300	3.01	–1.86	7.92
IFX	67,400	60,900	73,800	2.62	–2.54	7.73
RTX	69,400	62,600	76,200	3.10	–1.77	8.01
ABT	93,000	86,300	100,000	3.28	–1.52	8.02
DMARDs	49,000	43,300	55,100	2.13	–3.25	7.46

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	26,900	25,100	28,600	0.78	0.34	1.28
ETN–DMARDs	33,700	31,700	35,900	0.88	0.38	1.47
IFX–DMARDs	18,400	15,100	20,700	0.49	0.21	0.82
RTX–DMARDs	20,400	17,500	23,500	0.96	0.42	1.60
ABT–DMARDs	44,100	41,300	46,900	1.14	0.51	1.86
ADA–RTX	6,500	3,200	9,800	–0.18	–0.47	0.05
ETN–RTX	13,300	9,900	16,800	–0.09	–0.38	0.16
IFX–RTX	–2,000	–5,900	1,600	–0.48	–0.87	–0.16
ABT–RTX	23,600	19,600	27,500	0.18	–0.09	0.50
ADA–ABT	–17,200	–20,300	–14,100	–0.36	–0.72	–0.10
ETN–ABT	–10,400	–13,500	–7,100	–0.27	–0.59	–0.03
IFX–ABT	–25,600	–29,900	–22,100	–0.65	–1.12	–0.26
ADA–ETN	–6,800	–9,400	–4,200	–0.09	–0.32	0.10

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–IFX	8,400	5,700	12,000	0.29	0.07	0.56
ETN–IFX	15,300	12,300	18,900	0.39	0.14	0.72

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	34,300	20,900	79,000	0.01	0.31
ETN–DMARDs	38,400	23,200	87,400	0.00	0.18
IFX–DMARDs	37,700	22,100	90,300	0.01	0.20
RTX–DMARDs	21,200	12,800	48,400	0.39	0.84
ABT–DMARDs	38,500	23,400	86,600	0.00	0.17
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–TX	*4,100*	RTX	*16,000*	0.01	0.00
ABT–RTX	131,800	48,400	RTX	0.00	0.00
ADA–ABT	*47,700*	*23,500*	*177,100*	0.99	0.90
ETN–ABT	*38,900*	*16,300*	*308,500*	0.94	0.73
IFX–ABT	*39,100*	*22,400*	*95,800*	0.99	0.82
ADA–ETN	*72,800*	*20,400*	ADA	0.98	0.88
ADA–IFX	28,700	13,900	104,800	0.16	0.51
ETN–IFX	39,300	21,300	110,500	0.02	0.20

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective).

Same time on all biologics

In this scenario, the distribution of long-term survival time on all biologics was set to the value used for TNF inhibitors in the reference case. The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	74,800	68,700	81,000	2.88	–2.13	7.87
ETN	75,100	68,800	81,400	2.81	–2.26	7.84
IFX	73,000	66,000	79,900	2.80	–2.23	7.82
RTX	63,700	57,900	69,900	2.83	–2.15	7.86
ABT	82,000	75,700	88,600	2.97	–1.99	7.85
DMARDs	49,000	43,300	54,900	2.13	–3.23	7.46

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,800	24,100	27,600	0.75	0.33	1.21
ETN–DMARDs	26,100	24,400	27,900	0.68	0.28	1.12
IFX–DMARDs	24,000	19,300	26,800	0.67	0.29	1.12
RTX–DMARDs	14,700	13,600	15,900	0.70	0.30	1.15
ABT–DMARDs	33,000	30,800	35,400	0.84	0.37	1.37
ADA–RTX	11,100	9,200	13,100	0.05	–0.13	0.25
ETN–RTX	11,400	9,500	13,500	–0.02	–0.21	0.15
IFX–RTX	9,400	4,700	12,300	–0.03	–0.22	0.14
ABT–RTX	18,400	15,900	20,800	0.14	–0.05	0.36
ADA–ABT	–7,200	–10,000	–4,500	–0.09	–0.31	0.11
ETN–ABT	–6,900	–9,800	–4,000	–0.16	–0.40	0.03
IFX–ABT	–9,000	–14,100	–5,500	–0.17	–0.40	0.02
ADA–ETN	–300	–2,700	2,200	0.08	–0.10	0.28
ADA–IFX	1,800	–1,500	6,500	0.08	–0.10	0.29
ETN–IFX	2,100	–1,100	7,000	0.01	–0.17	0.18

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	34,400	20,900	78,400	0.01	0.31
ETN–DMARDs	38,700	23,300	91,700	0.01	0.17
IFX–DMARDs	35,900	21,200	81,100	0.02	0.24
RTX–DMARDs	21,100	12,600	49,100	0.41	0.84
ABT–DMARDs	39,500	23,800	89,700	0.00	0.15
ADA–RTX	206,000	44,700	RTX	0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	131,200	49,700	RTX	0.00	0.00
ADA–ABT	*84,100*	*22,500*	ADA	0.99	0.92
ETN–ABT	*42,700*	*15,900*	ETN	0.92	0.76
IFX–ABT	*53,700*	*20,800*	IFX	0.98	0.88
ADA–ETN	ADA	Not meaningful		0.82	0.82
ADA–IFX	21,600	Not meaningful		0.49	0.59
ETN–IFX	351,500	Not meaningful		0.19	0.25

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Rituximab cycle time 6 months

In this case, it was assumed that cycles of RTX would be given every 6 months. The assumption was that withdrawal rates per cycle would be maintained from the reference case. The results are as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	74,800	68,500	81,100	2.89	–2.16	7.81
ETN	75,100	69,000	81,500	2.80	–2.25	7.80
IFX	73,000	65,800	79,900	2.80	–2.27	7.81
RTX	74,800	67,200	82,400	2.93	–2.06	7.89
ABT	93,000	86,400	100,100	3.28	–1.52	8.05
DMARDs	49,000	43,400	55,000	2.13	–3.25	7.46

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,800	24,000	27,600	0.75	0.33	1.24
ETN–DMARDs	26,100	24,300	27,900	0.67	0.27	1.11
IFX–DMARDs	24,000	19,200	26,800	0.67	0.30	1.12
RTX–DMARDs	25,800	21,800	30,000	0.79	0.33	1.34
ABT–DMARDs	44,000	41,300	46,800	1.15	0.50	1.88
ADA–RTX	–18	–4,500	4,500	–0.04	–0.29	0.18
ETN–RTX	300	–4,100	4,700	–0.12	–0.38	0.09
IFX–RTX	–1,800	–7,500	3,200	–0.12	–0.38	0.08
ABT–RTX	18,200	13,200	23,100	0.35	0.07	0.73
ADA–ABT	–18,200	–21,300	–15,200	–0.39	–0.78	–0.12
ETN–ABT	–17,900	–21,100	–14,700	–0.47	–0.87	–0.18
IFX–ABT	–20,000	–25,400	–16,200	–0.48	–0.88	–0.16
ADA–ETN	–300	–2,800	2,100	0.08	–0.09	0.29
ADA–IFX	1,800	–1,500	6,500	0.08	–0.11	0.29
ETN–IFX	2,100	–1,200	7,200	0.00	–0.18	0.19

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	34,300	20,600	78,900	0.02	0.31
ETN–DMARDs	38,900	23,400	95,200	0.00	0.17
IFX–DMARDs	35,900	21,500	81,700	0.02	0.26
RTX–DMARDs	32,600	19,900	74,300	0.03	0.37
ABT–DMARDs	38,400	23,300	88,800	0.00	0.17
ADA–RTX	*430*	Not meaningful		0.36	0.35
ETN–RTX	RTX	Not meaningful		0.12	0.10
IFX–RTX	*14,700*	Not meaningful		0.40	0.28
ABT–RTX	51,500	25,400	229,200	0.00	0.07
ADA–ABT	*46,300*	*23,400*	*150,600*	1.00	0.90

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ETN–ABT	*37,800*	*20,300*	*95,700*	0.98	0.77
IFX–ABT	*42,000*	*22,500*	*117,700*	0.99	0.86
ADA–ETN	ADA	Not meaningful		0.83	0.84
ADA–IFX	21,700	Not meaningful		0.48	0.59
ETN–IFX	1,325,400	Not meaningful		0.18	0.23

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Rituximab cycle time 11.6 months

In this case, it was assumed that cycles of RTX would be given every 11.6 months, which was the observed mean time in the REFLEX extension study (Roche submission, p. 200). The assumption was that withdrawal rates per cycle would be maintained from the reference case. The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	74,800	68,600	81,200	2.89	–2.15	7.84
ETN	75,100	68,900	81,500	2.81	–2.23	7.87
IFX	73,000	66,000	80,000	2.80	–2.30	7.83
RTX	61,700	55,800	67,900	3.25	–1.58	8.11
ABT	93,100	86,100	100,100	3.28	–1.53	8.09
DMARDs	49,000	43,300	55,100	2.13	–3.27	7.49

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,800	24,100	27,600	0.75	0.32	1.24
ETN–DMARDs	26,100	24,300	28,000	0.67	0.29	1.10
IFX–DMARDs	24,000	19,300	26,800	0.67	0.28	1.10
RTX–DMARDs	12,700	11,000	14,500	1.11	0.49	1.81
ABT–DMARDs	44,000	41,300	46,800	1.15	0.52	1.89
ADA–RTX	13,100	10,800	15,500	–0.36	–0.74	–0.07
ETN–RTX	13,400	10,900	15,700	–0.44	–0.84	–0.14
IFX–RTX	11,300	6,600	14,600	–0.44	–0.85	–0.14
ABT–RTX	31,300	28,200	34,500	0.04	–0.27	0.33
ADA–ABT	–18,300	–21,400	–15,200	–0.39	–0.76	–0.11
ETN–ABT	–17,900	–21,100	–14,600	–0.48	–0.89	–0.17
IFX–ABT	–20,000	–25,400	–16,100	–0.48	–0.89	–0.17
ADA–ETN	–300	–2,700	2,100	0.08	–0.09	0.28
ADA–IFX	1,800	–1,600	6,300	0.09	–0.09	0.29
ETN–IFX	2,100	–1,400	6,900	0.00	–0.18	0.20

Diff, difference.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

Comparison	ICER	95% credible interval		Proportion of cases cost-effective at £20,000/QALY	£30,000/QALY
ADA–DMARDs	34,200	20,800	79,900	0.02	0.32
ETN–DMARDs	38,800	23,300	90,500	0.00	0.17
IFX–DMARDs	35,900	21,400	84,800	0.01	0.25
RTX–DMARDs	11,400	6,800	25,600	0.92	0.98
ABT–DMARDs	38,400	23,300	85,200	0.00	0.17
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	861,100	95,700	RTX	0.00	0.00
ADA–ABT	*46,400*	*23,600*	*150,400*	1.00	0.90
ETN–ABT	*37,800*	*20,100*	*103,800*	0.98	0.77
IFX–ABT	*41,800*	*22,600*	*120,500*	0.99	0.85
ADA–ETN	ADA	Not meaningful		0.83	0.84
ADA–IFX	20,700	Not meaningful		0.51	0.60
ETN–IFX	591,000	Not meaningful		0.18	0.24

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Poor late disease-modifying antirheumatic drugs (additional analysis)

In this scenario, the efficacy of conventional DMARDs taken after biologic therapy was reduced. HAQ multipliers were inferred from the Abbott and Roche industry submissions, and the lower of these figures (0.085) was taken. Preserving $a+b=1.22$ from the BRAM reference case for LEF gave $a=0.104$, $b=1.116$. These values were then used for all conventional DMARDs. The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	76,400	69,600	83,100	2.30	–3.02	7.59
ETN	76,600	70,000	83,300	2.23	–3.06	7.55
IFX	74,600	67,100	81,800	2.22	–3.12	7.51
RTX	70,700	63,800	78,000	2.61	–2.58	7.72
ABT	94,400	87,300	101,800	2.76	–2.33	7.81
DMARDs	51,000	44,900	57,300	1.40	–4.38	7.15

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,400	23,700	27,200	0.90	0.41	1.47
ETN–DMARDs	25,600	23,900	27,400	0.82	0.36	1.33
IFX–DMARDs	23,600	19,000	26,400	0.82	0.36	1.35
RTX–DMARDs	19,700	16,900	22,400	1.21	0.53	1.95
ABT–DMARDs	43,400	40,700	46,200	1.35	0.62	2.17
ADA–RTX	5,700	2,500	8,800	–0.30	–0.66	–0.03
ETN–RTX	6,000	2,800	9,100	–0.38	–0.77	–0.08

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
IFX–RTX	3,900	–1,100	7,800	–0.39	–0.77	–0.09
ABT–RTX	23,700	19,900	27,500	0.15	–0.14	0.48
ADA–ABT	–18,100	–21,300	–15,000	–0.45	–0.83	–0.17
ETN–ABT	–17,800	–21,000	–14,600	–0.53	–0.96	–0.21
IFX–ABT	–19,800	–25,100	–16,000	–0.54	–0.94	–0.20
ADA–ETN	–300	–2,600	2,100	0.08	–0.10	0.28
ADA–IFX	1,800	–1,500	6,500	0.09	–0.09	0.29
ETN–IFX	2,000	–1,300	7,000	0.01	–0.17	0.20

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	28,100	17,200	62,300	0.10	0.57
ETN–DMARDs	31,100	19,200	70,400	0.04	0.43
IFX–DMARDs	28,800	17,200	63,400	0.08	0.54
RTX–DMARDs	16,300	10,100	36,100	0.73	0.94
ABT–DMARDs	32,100	20,000	71,600	0.02	0.39
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	158,600	51,500	RTX	0.00	0.00
ADA–ABT	*40,100*	*21,400*	*106,200*	0.99	0.82
ETN–ABT	*33,500*	*18,400*	*81,400*	0.95	0.67
IFX–ABT	*36,900*	*20,500*	*95,800*	0.98	0.76
ADA–ETN	ADA	Not meaningful		0.83	0.84
ADA–IFX	20,600	Not meaningful		0.50	0.61
ETN–IFX	316,000	Not meaningful		0.17	0.23

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Health Assessment Questionnaire change on biologics

In this scenario, a deterioration of 0.03/year in HAQ was assumed on biologic treatments. This was modelled as a mean time between 0.125-unit increases of 4 years. For each treatment separately, this figure was given a normal distribution with a SD of 0.4 years. The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	75,500	69,200	81,900	2.53	–2.56	7.75
ETN	75,800	69,500	82,200	2.44	–2.69	7.74
IFX	73,700	66,500	80,700	2.45	–2.71	7.73
RTX	70,400	63,600	77,500	2.56	–2.50	7.81
ABT	93,900	86,900	101,200	2.80	–2.17	7.91
DMARDs	49,100	43,500	54,700	2.09	–3.17	7.50

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	26,400	24,600	28,400	0.43	0.16	0.78
ETN–DMARDs	26,700	24,800	28,700	0.35	0.12	0.64
IFX–DMARDs	24,600	20,300	27,300	0.36	0.12	0.66
RTX–DMARDs	21,300	18,200	24,400	0.46	0.15	0.85
ABT–DMARDs	44,800	42,000	47,700	0.71	0.29	1.22
ADA–RTX	5,100	1,600	8,500	–0.03	–0.30	0.20
ETN–RTX	5,400	2,000	8,700	–0.11	–0.41	0.11
IFX–RTX	3,300	–2,000	7,400	–0.11	–0.39	0.12
ABT–RTX	23,500	19,600	27,400	0.24	–0.03	0.61
ADA–ABT	–18,400	–21,400	–15,300	–0.28	–0.59	–0.02
ETN–ABT	–18,100	–21,100	–15,000	–0.36	–0.72	–0.09
IFX–ABT	–20,100	–25,000	–16,300	–0.35	–0.72	–0.08
ADA–ETN	–300	–2,700	2,100	0.08	–0.11	0.32
ADA–IFX	1,800	–1,500	6,300	0.07	–0.12	0.29
ETN–IFX	2,100	–1,300	6,700	–0.01	–0.21	0.18

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	61,300	33,600	168,600	0.00	0.01
ETN–DMARDs	76,300	42,500	228,200	0.00	0.00
IFX–DMARDs	68,900	36,200	200,000	0.00	0.00
RTX–DMARDs	46,000	24,600	134,400	0.00	0.09
ABT–DMARDs	63,300	36,700	151,700	0.00	0.00
ADA–RTX	RTX	Not meaningful		0.02	0.06
ETN–RTX	RTX	Not meaningful		0.00	0.01
IFX–RTX	RTX	Not meaningful		0.05	0.05
ABT–RTX	96,400	38,900	RTX	0.00	0.00

Comparison	ICER	95% credible interval		Proportion of cases cost-effective at £20,000/QALY	£30,000/QALY
ADA–ABT	*66,500*	*29,400*	*718,000*	1.00	0.97
ETN–ABT	*50,600*	*24,300*	*205,800*	0.99	0.91
IFX–ABT	*57,600*	*27,400*	*250,500*	1.00	0.96
ADA–ETN	ADA	Not meaningful		0.78	0.80
ADA–IFX	24,300	Not meaningful		0.46	0.55
ETN–IFX	IFX	Not meaningful		0.21	0.26

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Adverse event costs included

Additional annual costs based on the Bristol-Myers Squibb LTD submission as follows:

Treatment	Additional cost (£)
ADA	117.82
ETN	224.87
IFX	162.02
RTX	273.51
ABT	110.16

When these were included, the results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	75,100	69,200	81,400	2.89	–2.12	7.87
ETN	75,700	69,400	82,200	2.80	–2.21	7.84
IFX	73,500	66,500	80,300	2.80	–2.24	7.82
RTX	70,700	63,800	77,700	3.10	–1.78	7.95
ABT	93,500	86,600	100,600	3.28	–1.46	8.05
DMARDs	49,000	43,300	54,900	2.13	–3.27	7.46

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	26,100	24,500	27,900	0.75	0.33	1.23
ETN–DMARDs	26,800	24,900	28,700	0.67	0.30	1.10
IFX–DMARDs	24,500	19,800	27,300	0.67	0.29	1.12
RTX–DMARDs	21,700	18,600	24,700	0.96	0.41	1.61
ABT–DMARDs	44,500	41,700	47,200	1.15	0.52	1.88
ADA–RTX	4,500	1,200	8,000	–0.21	–0.52	0.03
ETN–RTX	5,100	1,600	8,700	–0.29	–0.63	–0.04
IFX–RTX	2,800	–2,500	7,000	–0.30	–0.62	–0.05
ABT–RTX	22,800	18,800	26,800	0.18	–0.10	0.50
ADA–ABT	–18,300	–21,500	–15,300	–0.39	–0.77	–0.12
ETN–ABT	–17,700	–21,100	–14,300	–0.47	–0.88	–0.17

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
IFX–ABT	–20,000	–25,200	–16,100	–0.48	–0.88	–0.17
ADA–ETN	–600	–3,200	1,800	0.08	–0.09	0.29
ADA–IFX	1,600	–1,600	6,400	0.09	–0.10	0.29
ETN–IFX	2,200	–1,100	7,100	0.00	–0.17	0.19

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	34,700	21,200	80,200	0.01	0.29
ETN–DMARDs	39,900	24,200	91,400	0.00	0.14
IFX–DMARDs	36,800	21,700	83,700	0.01	0.22
RTX–DMARDs	22,500	13,700	52,800	0.32	0.80
ABT–DMARDs	38,800	23,300	85,600	0.00	0.17
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	126,100	46300	RTX	0.00	0.00
ADA–ABT	*46,700*	*23,200*	*153,000*	0.99	0.90
ETN–ABT	*37,400*	*19,800*	*101,100*	0.97	0.76
IFX–ABT	*41,700*	*21,900*	*113,300*	0.99	0.84
ADA–ETN	ADA	Not meaningful		0.87	0.87
ADA–IFX	19,000	Not meaningful		0.53	0.63
ETN–IFX	502,600	Not meaningful		0.17	0.22

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

No offset costs (additional analysis)

In this case the 'offset costs' representing the estimates of joint replacement and hospitalisation costs were removed. The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	47,200	44,500	49,800	2.89	–2.12	7.87
ETN	47,200	44,400	50,000	2.80	–2.21	7.84
IFX	45,100	40,100	48,700	2.80	–2.24	7.82
RTX	42,100	38,200	46,100	3.10	–1.78	7.95
ABT	66,400	62,500	70,400	3.28	–1.46	8.05
DMARDs	19,400	17,900	20,900	2.13	–3.27	7.46

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	27,800	26,000	29,600	0.75	0.33	1.23
ETN–DMARDs	27,800	25,800	29,800	0.67	0.30	1.10
IFX–DMARDs	25,700	21,000	28,500	0.67	0.29	1.12
RTX–DMARDs	22,700	19,500	26,000	0.96	0.41	1.61
ABT–DMARDs	47,000	44,000	50,100	1.15	0.52	1.88
ADA–RTX	5,000	1,600	8,500	–0.21	–0.52	0.03
ETN–RTX	5,100	1,500	8,800	–0.29	–0.63	–0.04
IFX–RTX	3,000	–2,400	7,100	–0.30	–0.62	–0.05
ABT–RTX	24,300	20,200	28,400	0.18	–0.10	0.50
ADA–ABT	–19,200	–22,400	–16,200	–0.39	–0.77	–0.12
ETN–ABT	–19,200	–22,500	–16,000	–0.47	–0.88	–0.17
IFX–ABT	–21,300	–26,500	–17,400	–0.48	–0.88	–0.17
ADA–ETN	–33	–2,400	2,400	0.08	–0.09	0.29
ADA–IFX	2,000	–1,200	6,700	0.09	–0.10	0.29
ETN–IFX	2,100	–1,300	6,900	0.00	–0.17	0.19

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	36,900	22,800	84,200	0.01	0.22
ETN–DMARDs	41,400	25,400	95,100	0.00	0.11
IFX–DMARDs	38,600	23,100	89,600	0.01	0.17
RTX–DMARDs	23,600	14,600	55,300	0.27	0.76
ABT–DMARDs	41,000	24,900	90,700	0.00	0.11
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	134,100	50,100	RTX	0.00	0.00
ADA–ABT	*49,000*	*25,100*	*153,200*	1.00	0.92
ETN–ABT	*40,500*	*22,100*	*109,500*	0.99	0.83
IFX–ABT	*44,400*	*24,000*	*118,000*	1.00	0.89
ADA–ETN	ADA	Not meaningful		0.83	0.84
ADA–IFX	23,500	Not meaningful		0.46	0.59
ETN–IFX	460,000	Not meaningful		0.17	0.23

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Extra cost for palliation (additional analysis)

In this scenario, the cost for Pall was increased to the cost of methotrexate, including monitoring. This involved a start-up cost of £421.03 and an annual usage cost of £995.78. The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	76,800	70,500	83,500	2.89	–2.12	7.87
ETN	77,100	70,300	84,000	2.80	–2.21	7.84
IFX	75,000	67,800	82,200	2.80	–2.24	7.82
RTX	71,100	64,100	78,300	3.10	–1.78	7.95
ABT	94,800	87,700	102,300	3.28	–1.46	8.05
DMARDs	51,700	45,500	58,300	2.13	–3.27	7.46

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,100	23,400	26,800	0.75	0.33	1.23
ETN–DMARDs	25,400	23,600	27,200	0.67	0.30	1.10
IFX–DMARDs	23,400	18,900	26,100	0.67	0.29	1.12
RTX–DMARDs	19,400	16,600	22,200	0.96	0.41	1.61
ABT–DMARDs	43,100	40,400	45,700	1.15	0.52	1.88
ADA–RTX	5,700	2,600	8,900	–0.21	–0.52	0.03
ETN–RTX	6,000	2,800	9,300	–0.29	–0.63	–0.04
IFX–RTX	4,000	–1,100	7,800	–0.30	–0.62	–0.05
ABT–RTX	23,700	19,900	27,400	0.18	–0.10	0.50
ADA–ABT	–18,000	–21,000	–15,000	–0.39	–0.77	–0.12
ETN–ABT	–17,700	–20,900	–14,400	–0.47	–0.88	–0.17
IFX–ABT	–19,700	–24,700	–16,100	–0.48	–0.88	–0.17
ADA–ETN	–300	–2,700	2,000	0.08	–0.09	0.29
ADA–IFX	1,700	–1,400	6,300	0.09	–0.10	0.29
ETN–IFX	2,000	–1,200	6,700	0.00	–0.17	0.19

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	33,400	20,400	76,600	0.02	0.34
ETN–DMARDs	37,800	22,900	86,800	0.01	0.20
IFX–DMARDs	35,000	20,600	79,600	0.02	0.28
RTX–DMARDs	20,100	12,100	47,400	0.46	0.86
ABT–DMARDs	37,600	22,600	83,000	0.01	0.19
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	131,000	47,800	RTX	0.00	0.00
ADA–ABT	*45,800*	*22,800*	*150,000*	0.99	0.89

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ETN–ABT	*37,300*	*19,800*	*100,800*	0.97	0.76
IFX–ABT	*41,200*	*21,700*	*112,800*	0.99	0.83
ADA–ETN	ADA	Not meaningful		0.84	0.84
ADA–IFX	20,300	Not meaningful		0.50	0.61
ETN–IFX	452,000	Not meaningful		0.19	0.25

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

No negative quality of life scores

In this case, all QoL scores that were calculated as negative using the equation converting HAQ to QoL were replaced by zero. The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	74,800	68,800	81,000	3.80	1.68	7.87
ETN	75,100	68,700	81,500	3.73	1.64	7.84
IFX	73,000	66,100	79,700	3.73	1.66	7.82
RTX	69,400	62,700	76,400	3.94	1.77	7.95
ABT	93,000	86,200	100,100	4.11	1.97	8.05
DMARDs	49,000	43,300	54,900	3.27	1.32	7.46

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,800	24,100	27,500	0.53	0.29	0.73
ETN–DMARDs	26,100	24,200	27,900	0.46	0.25	0.66
IFX–DMARDs	24,000	19,500	26,800	0.46	0.24	0.66
RTX–DMARDs	20,400	17,500	23,200	0.67	0.35	0.95
ABT–DMARDs	44,000	41,300	46,700	0.83	0.50	1.12
ADA–RTX	5,400	2,200	8,700	–0.13	–0.36	0.07
ETN–RTX	5,700	2,400	9,100	–0.20	–0.43	0.00
IFX–RTX	3,600	–1,600	7,600	–0.20	–0.43	0.00
ABT–RTX	23,600	19,800	27,400	0.17	–0.08	0.42
ADA–ABT	–18,200	–21,300	–15,200	–0.30	–0.54	–0.08
ETN–ABT	–18,000	–21,200	–14,600	–0.37	–0.61	–0.15
IFX–ABT	–20,000	–25,100	–16,200	–0.37	–0.60	–0.15
ADA–ETN	–300	–2,800	2,100	0.07	–0.08	0.23
ADA–IFX	1,800	–1,400	6,500	0.07	–0.10	0.23
ETN–IFX	2,000	–1,200	6,800	0.00	–0.15	0.15

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

Comparison	ICER	95% credible interval		Proportion of cases cost-effective at £20,000/QALY	Proportion of cases cost-effective at £30,000/QALY
ADA–DMARDs	48,600	35,300	87,100	0.00	0.00
ETN–DMARDs	56,500	39,100	102,700	0.00	0.00
IFX–DMARDs	52,100	35,100	97,500	0.00	0.00
RTX–DMARDs	30,700	21,700	57,300	0.01	0.48
ABT–DMARDs	52,800	39,100	89,100	0.00	0.00
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	140,700	58,000	RTX	0.00	0.00
ADA–ABT	*60,300*	*33,200*	*216,900*	1.00	0.99
ETN–ABT	*48,300*	*29,100*	*115,000*	1.00	0.97
IFX–ABT	*53,700*	*31,700*	*129,100*	1.00	0.99
ADA–ETN	ADA	Not meaningful		0.82	0.84
ADA–IFX	25,300	Not meaningful		0.46	0.57
ETN–IFX	7,430,000	Not meaningful		0.18	0.22

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Linear equation Health Assessment Questionnaire to quality of life

In this scenario, the linear equation QoL = 0.862 – 0.327HAQ was used as in previous versions of the BRAM. For the probabilistic analysis, the coefficients were sampled from normal distributions with SDs 0.034 and 0.0201 respectively.[174] The results were as follows:

Treatment	Mean cost	95% credible interval		Mean QALY	95% credible interval	
ADA	74,800	68,700	80,900	3.03	1.66	4.35
ETN	75,100	68,900	81,600	2.96	1.59	4.28
IFX	73,000	66,000	79,800	2.95	1.60	4.28
RTX	69,400	62,700	76,000	3.22	1.88	4.55
ABT	93,000	86,300	99,600	3.40	2.05	4.71
DMARDs	49,000	43,300	54,900	2.36	0.97	3.72

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–DMARDs	25,800	24,100	27,600	0.67	0.51	0.84
ETN–DMARDs	26,100	24,300	28,000	0.60	0.44	0.76
IFX–DMARDs	24,100	19,600	26,800	0.59	0.44	0.76
RTX–DMARDs	20,400	17,800	23,300	0.86	0.63	1.12
ABT–DMARDs	44,000	41,400	46,600	1.04	0.81	1.29
ADA–RTX	5,400	2,000	8,600	–0.19	–0.43	0.04
ETN–RTX	5,700	2,600	9,000	–0.26	–0.50	–0.04
IFX–RTX	3,700	–1,500	7,600	–0.27	–0.53	–0.04
ABT–RTX	23,700	19,700	27,400	0.18	–0.10	0.44

Comparison	Diff cost	95% credible interval		Diff QALY	95% credible interval	
ADA–ABT	–18,200	–21,300	–15,100	–0.37	–0.61	–0.15
ETN–ABT	–17,900	–21,100	–14,800	–0.44	–0.67	–0.23
IFX–ABT	–20,000	–25,200	–16,200	–0.45	–0.70	–0.22
ADA–ETN	–300	–2,800	2,000	0.07	–0.10	0.25
ADA–IFX	1,800	–1,400	6,400	0.08	–0.11	0.26
ETN–IFX	2,100	–1,100	6,700	0.00	–0.17	0.17

Diff, difference in.
Difference calculated by subtracting the value for the strategy named second from the value for the strategy named first.

				Proportion of cases cost-effective at	
Comparison	ICER	95% credible interval		£20,000/QALY	£30,000/QALY
ADA–DMARDs	38,600	30,300	51,000	0.00	0.02
ETN–DMARDs	43,800	34,600	57,900	0.00	0.00
IFX–DMARDs	40,600	30,700	54,100	0.00	0.02
RTX–DMARDs	23,700	18,700	31,100	0.08	0.96
ABT–DMARDs	42,300	34,100	54,100	0.00	0.00
ADA–RTX	RTX	Not meaningful		0.00	0.00
ETN–RTX	RTX	Not meaningful		0.00	0.00
IFX–RTX	RTX	Not meaningful		0.00	0.00
ABT–RTX	130,900	54,400	RTX	0.00	0.00
ADA–ABT	*49,100*	*29,300*	*120,000*	1.00	0.97
ETN–ABT	*40,300*	*26,100*	*77,500*	1.00	0.90
IFX–ABT	*44,600*	*27,900*	*92,400*	1.00	0.95
ADA–ETN	ADA	Not meaningful		0.81	0.82
ADA–IFX	23,100	Not meaningful		0.48	0.57
ETN–IFX	667,000	Not meaningful		0.19	0.25

The proportion of cases cost-effective relates to the strategy given first on each line. ICER in italics means that the strategy named second is more costly and more effective. Where a strategy name is given in place of an ICER, the named strategy dominates its comparator (less costly and more effective). A 95% credible interval for the ICER is not meaningful in cases where the cost-effectiveness scatter plot is not confined to one half of the plane.

Appendix 15

Assumptions in the Birmingham Rheumatoid Arthritis Model

Item	Data source or assumption	Comments
Baseline characteristics	Based on the British Society for Rheumatology (BSR) submission to the NICE	Most recent UK data reflecting characteristics of patients in actual clinical practice
Sequence of treatments	Analysed strategies after the failure of a tumour necrosis factor (TNF) inhibitor are: **ADA** / **ETN** / **IFX** / **RTX** / **ABT** / **DMARD** ADA / ETN / IFX / RTX / ABT / DMARD LEF / LEF / LEF / LEF / LEF / GST GST / GST / GST / GST / GST / CyA CyA / CyA / CyA / CyA / CyA / AZA AZA / AZA / AZA / AZA / AZA / Pall Pall / Pall / Pall / Pall / Pall DMARDs that are likely to have been used before biologic therapy are excluded from the sequence	We restricted the analysis to a second biologic therapy and did not consider sequences including a third biologic therapy. We assumed that the effectiveness of the 'late' DMARDs (LEF, GST, CyA, AZA) would not be as good as when these drugs were used in early RA, but would be equally good between the strategies modelled
Health Assessment Questionnaire (HAQ) change on initiation of treatment	Biologics – data from randomised controlled trials and if not available from largest observational cohorts studies; we have assumed that IFX was equally effective to ETN DMARDs – no data for patients who have failed a TNF inhibitor were available; assumed to be half of the HAQ improvement in early RA. However, much lower responses are explored in sensitivity analyses	The best available evidence was used
Health Assessment Questionnaire change on treatment	It is assumed that after initial improvement the HAQ score changes on treatment by: ■ 0 on biologic treatments ■ 0.045/year on conventional DMARD ■ 0.06/year on Pall This is modelled as mean times to an increase in HAQ score of 0.125 (2.7 for DMARDs and 2.0 for Pall) In PSA times are sampled from normal distributions with SD of 0.27 (DMARD) and 0.2 (Pall)	These assumptions are in line with those made in previous versions of the BRAM. However, it should be appreciated that even with the optimal treatment a majority of patients do not achieve remission. Therefore, because continuing disease activity is likely to have a detrimental effect on physical function, an assumption of zero HAQ progression on biologic treatments in models that span a lifetime is somewhat implausible
Health Assessment Questionnaire increase on withdrawing from treatment	It is assumed to be the same as initial improvement on treatment	This has been an assumption in all versions of the BRAM and in other models
Time on treatment – probability of early quitting	Data from the following sources is used: ■ DMARDs – from Chen *et al.*[179] ■ ADA – from Bombardieri *et al.*[96] ■ ETN – from Bingham *et al.*[104] and Buch *et al.*[99] ■ IFX – from OPPOSITE[133] ■ RTX – not possible in the model ■ ABT – from ATTAIN[127–132]	The best available evidence was used. For biologic drugs the highest quality sources identified in the systematic review were used. For DMARDs no studies where DMARDs were used after failure of a TNF inhibitor were identified. Therefore, data from early RA were used. We are aware that disease duration can influence HAQ responses and taking values from studies in early RA is problematic. However, halving a HAQ response, for example for leflunomide from 0.38 to 0.19 (approaching the minimal clinical detectable difference), is plausible

Item	Data source or assumption	Comments
Time to withdrawal – long-term survival on treatment	For long-term survival on treatment Weibull curves were fitted to the available data: ■ TNF inhibitors – from BSR submission[163] ■ RTX – from REFLEX LTE[139] ■ ABT – from BMS submission[205] ■ DMARDS – General Practice Research Database (GPRD) data	The best available evidence was used. For TNF inhibitors most recent UK data were chosen and as it was not available for RTX and ABT data from clinical trials were utilised. For DMARDs no studies in the relevant population were identified and therefore data from GPRD were used
Mortality	Basic mortality was taken from standard life tables. A RR (1.33) per unit HAQ was applied. For PSA a log-normal distribution was assumed (95% CI 1.10 to 1.61)	Based on Wolfe *et al.*[199] Previous versions of the BRAM were not found to be sensitive to this parameter
QoL scores	The following equation was used to map HAQ onto QoL: $QoL = a - b_1 \times HAQ - b_2 \times HAQ^2$ This allows negative utility values A scenario analysis that adjusted all negative utilities to zero is reported	We used the results of a regression performed on data from Hurst *et al.*[155] in the absence of any more recent data
Costs	Costs are made up of drug and monitoring costs. A 'start-up' cost reflects higher dosage and additional monitoring, as appropriate for each treatment Unit costs were based on: ■ For tests and visits – values from Chen *et al.*[179] inflated to 2008 and from Curtis[191] ■ For drugs – *British National Formulary* 58 accessed online	This simplifying assumption means that all patients incur the full additional 'start-up' costs even if quitting early. In most cases, the additional costs are complete within 3 months of starting; only in the case of GST do the additional costs extend beyond 6 months
Monitoring assumptions	The data on monitoring was: ■ For DMARDs – based on Chen *et al.*[179] ■ For biologics – assumed to be the same as for MTX – and all drugs are given with concomitant MTX	It was assumed that there will be no difference in the monitoring necessary for DMARDs between early and late RA and therefore data from Chen *et al.*[179] were used. For biologics it was assumed that all of them are given with MTX and therefore MTX monitoring was assumed to be sufficient

Health Technology Assessment programme

HTA Commissioning Board *(continued)*

Professor James Raftery,
Chair of NETSCC and Director of the Wessex Institute, University of Southampton

Professor Barney Reeves,
Professorial Research Fellow in Health Services Research, Department of Clinical Science, University of Bristol

Professor Martin Underwood,
Warwick Medical School, University of Warwick

Professor Marion Walker,
Professor in Stroke Rehabilitation, Associate Director UK Stroke Research Network, University of Nottingham

Dr Duncan Young,
Senior Clinical Lecturer and Consultant, Nuffield Department of Anaesthetics, University of Oxford

Observers

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

HTA Clinical Evaluation and Trials Board

Chair,
Professor Sallie Lamb,
Director,
Warwick Clinical Trials Unit,
Warwick Medical School,
University of Warwick and Professor of Rehabilitation,
Nuffield Department of Orthopaedic, Rheumatology and Musculoskeletal Sciences,
University of Oxford

Deputy Chair,
Professor Jenny Hewison,
Professor of the Psychology of Health Care,
Leeds Institute of Health Sciences,
University of Leeds

Programme Director,
Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Members

Professor Keith Abrams,
Professor of Medical Statistics, Department of Health Sciences, University of Leicester

Professor Martin Bland,
Professor of Health Statistics, Department of Health Sciences, University of York

Professor Jane Blazeby,
Professor of Surgery and Consultant Upper GI Surgeon, Department of Social Medicine, University of Bristol

Professor Julia M Brown,
Director, Clinical Trials Research Unit, University of Leeds

Professor Alistair Burns,
Professor of Old Age Psychiatry, Psychiatry Research Group, School of Community-Based Medicine, The University of Manchester & National Clinical Director for Dementia, Department of Health

Dr Jennifer Burr,
Director, Centre for Healthcare Randomised trials (CHART), University of Aberdeen

Professor Linda Davies,
Professor of Health Economics, Health Sciences Research Group, University of Manchester

Professor Simon Gilbody,
Prof of Psych Medicine and Health Services Research, Department of Health Sciences, University of York

Professor Steven Goodacre,
Professor and Consultant in Emergency Medicine, School of Health and Related Research, University of Sheffield

Professor Dyfrig Hughes,
Professor of Pharmacoeconomics, Centre for Economics and Policy in Health, Institute of Medical and Social Care Research, Bangor University

Professor Paul Jones,
Professor of Respiratory Medicine, Department of Cardiac and Vascular Science, St George's Hospital Medical School, University of London

Professor Khalid Khan,
Professor of Women's Health and Clinical Epidemiology, Barts and the London School of Medicine, Queen Mary, University of London

Professor Richard J McManus,
Professor of Primary Care Cardiovascular Research, Primary Care Clinical Sciences Building, University of Birmingham

Professor Helen Rodgers,
Professor of Stroke Care, Institute for Ageing and Health, Newcastle University

Professor Ken Stein,
Professor of Public Health, Peninsula Technology Assessment Group, Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth

Professor Jonathan Sterne,
Professor of Medical Statistics and Epidemiology, Department of Social Medicine, University of Bristol

Mr Andy Vail,
Senior Lecturer, Health Sciences Research Group, University of Manchester

Professor Clare Wilkinson,
Professor of General Practice and Director of Research North Wales Clinical School, Department of Primary Care and Public Health, Cardiff University

Dr Ian B Wilkinson,
Senior Lecturer and Honorary Consultant, Clinical Pharmacology Unit, Department of Medicine, University of Cambridge

Observers

Ms Kate Law,
Director of Clinical Trials, Cancer Research UK

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Diagnostic Technologies and Screening Panel

Members

Chair,
Professor Lindsay Wilson Turnbull,
Scientific Director of the Centre for Magnetic Resonance Investigations and YCR Professor of Radiology, Hull Royal Infirmary

Professor Judith E Adams,
Consultant Radiologist, Manchester Royal Infirmary, Central Manchester & Manchester Children's University Hospitals NHS Trust, and Professor of Diagnostic Radiology, University of Manchester

Mr Angus S Arunkalaivanan,
Honorary Senior Lecturer, University of Birmingham and Consultant Urogynaecologist and Obstetrician, City Hospital, Birmingham

Dr Stephanie Dancer,
Consultant Microbiologist, Hairmyres Hospital, East Kilbride

Dr Diane Eccles,
Professor of Cancer Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital

Dr Trevor Friedman,
Consultant Liason Psychiatrist, Brandon Unit, Leicester General Hospital

Dr Ron Gray,
Consultant, National Perinatal Epidemiology Unit, Institute of Health Sciences, University of Oxford

Professor Paul D Griffiths,
Professor of Radiology, Academic Unit of Radiology, University of Sheffield

Mr Martin Hooper,
Public contributor

Professor Anthony Robert Kendrick,
Associate Dean for Clinical Research and Professor of Primary Medical Care, University of Southampton

Dr Anne Mackie,
Director of Programmes, UK National Screening Committee, London

Mr David Mathew,
Public contributor

Dr Michael Millar,
Consultant Senior Lecturer in Microbiology, Department of Pathology & Microbiology, Barts and The London NHS Trust, Royal London Hospital

Mrs Una Rennard,
Public contributor

Dr Stuart Smellie,
Consultant in Clinical Pathology, Bishop Auckland General Hospital

Ms Jane Smith,
Consultant Ultrasound Practitioner, Leeds Teaching Hospital NHS Trust, Leeds

Dr Allison Streetly,
Programme Director, NHS Sickle Cell and Thalassaemia Screening Programme, King's College School of Medicine

Dr Alan J Williams,
Consultant Physician, General and Respiratory Medicine, The Royal Bournemouth Hospital

Observers

Dr Tim Elliott,
Team Leader, Cancer Screening, Department of Health

Dr Catherine Moody,
Programme Manager, Medical Research Council

Professor Julietta Patrick,
Director, NHS Cancer Screening Programme, Sheffield

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Disease Prevention Panel

Members

Chair,
Professor Margaret Thorogood,
Professor of Epidemiology, University of Warwick Medical School, Coventry

Dr Robert Cook,
Clinical Programmes Director, Bazian Ltd, London

Dr Colin Greaves,
Senior Research Fellow, Peninsula Medical School (Primary Care)

Mr Michael Head,
Public contributor

Professor Cathy Jackson,
Professor of Primary Care Medicine, Bute Medical School, University of St Andrews

Dr Russell Jago,
Senior Lecturer in Exercise, Nutrition and Health, Centre for Sport, Exercise and Health, University of Bristol

Dr Julie Mytton,
Consultant in Child Public Health, NHS Bristol

Professor Irwin Nazareth,
Professor of Primary Care and Director, Department of Primary Care and Population Sciences, University College London

Dr Richard Richards,
Assistant Director of Public Health, Derbyshire Country Primary Care Trust

Professor Ian Roberts,
Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine

Dr Kenneth Robertson,
Consultant Paediatrician, Royal Hospital for Sick Children, Glasgow

Dr Catherine Swann,
Associate Director, Centre for Public Health Excellence, NICE

Professor Carol Tannahill,
Glasgow Centre for Population Health

Mrs Jean Thurston,
Public contributor

Professor David Weller,
Head, School of Clinical Science and Community Health, University of Edinburgh

Observers

Ms Christine McGuire,
Research & Development, Department of Health

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

External Devices and Physical Therapies Panel

Members

Chair,
Dr John Pounsford,
Consultant Physician North Bristol NHS Trust

Deputy Chair,
Professor E Andrea Nelson,
Reader in Wound Healing and Director of Research, University of Leeds

Professor Bipin Bhakta,
Charterhouse Professor in Rehabilitation Medicine, University of Leeds

Mrs Penny Calder,
Public contributor

Dr Dawn Carnes,
Senior Research Fellow, Barts and the London School of Medicine and Dentistry

Dr Emma Clark,
Clinician Scientist Fellow & Cons. Rheumatologist, University of Bristol

Mrs Anthea De Barton-Watson,
Public contributor

Professor Nadine Foster,
Professor of Musculoskeletal Health in Primary Care Arthritis Research, Keele University

Dr Shaheen Hamdy,
Clinical Senior Lecturer and Consultant Physician, University of Manchester

Professor Christine Norton,
Professor of Clinical Nursing Innovation, Bucks New University and Imperial College Healthcare NHS Trust

Dr Lorraine Pinnigton,
Associate Professor in Rehabilitation, University of Nottingham

Dr Kate Radford,
Senior Lecturer (Research), University of Central Lancashire

Mr Jim Reece,
Public contributor

Professor Maria Stokes,
Professor of Neuromusculoskeletal Rehabilitation, University of Southampton

Dr Pippa Tyrrell,
Senior Lecturer/Consultant, Salford Royal Foundation Hospitals' Trust and University of Manchester

Dr Sarah Tyson,
Senior Research Fellow & Associate Head of School, University of Salford

Dr Nefyn Williams,
Clinical Senior Lecturer, Cardiff University

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Interventional Procedures Panel

Members

Chair,
Professor Jonathan Michaels,
Professor of Vascular Surgery, University of Sheffield

Deputy Chair,
Mr Michael Thomas,
Consultant Colorectal Surgeon, Bristol Royal Infirmary

Mrs Isabel Boyer,
Public contributor

Mr David P Britt,
Public contributor

Mr Sankaran Chandra Sekharan,
Consultant Surgeon, Breast Surgery, Colchester Hospital University NHS Foundation Trust

Professor Nicholas Clarke,
Consultant Orthopaedic Surgeon, Southampton University Hospitals NHS Trust

Ms Leonie Cooke,
Public contributor

Mr Seumas Eckford,
Consultant in Obstetrics & Gynaecology, North Devon District Hospital

Professor Sam Eljamel,
Consultant Neurosurgeon, Ninewells Hospital and Medical School, Dundee

Dr Adele Fielding,
Senior Lecturer and Honorary Consultant in Haematology, University College London Medical School

Dr Matthew Hatton,
Consultant in Clinical Oncology, Sheffield Teaching Hospital Foundation Trust

Dr John Holden,
General Practitioner, Garswood Surgery, Wigan

Professor Nicholas James,
Professor of Clinical Oncology, School of Cancer Sciences, University of Birmingham

Dr Fiona Lecky,
Senior Lecturer/Honorary Consultant in Emergency Medicine, University of Manchester/Salford Royal Hospitals NHS Foundation Trust

Dr Nadim Malik,
Consultant Cardiologist/Honorary Lecturer, University of Manchester

Mr Hisham Mehanna,
Consultant & Honorary Associate Professor, University Hospitals Coventry & Warwickshire NHS Trust

Dr Jane Montgomery,
Consultant in Anaesthetics and Critical Care, South Devon Healthcare NHS Foundation Trust

Professor Jon Moss,
Consultant Interventional Radiologist, North Glasgow Hospitals University NHS Trust

Dr Simon Padley,
Consultant Radiologist, Chelsea & Westminster Hospital

Dr Ashish Paul,
Medical Director, Bedfordshire PCT

Dr Sarah Purdy,
Consultant Senior Lecturer, University of Bristol

Professor Yit Chiun Yang,
Consultant Ophthalmologist, Royal Wolverhampton Hospitals NHS Trust

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Pharmaceuticals Panel

Members

Chair,
Professor Imti Choonara,
Professor in Child Health, University of Nottingham

Deputy Chair,
Dr Yoon K Loke,
Senior Lecturer in Clinical Pharmacology, University of East Anglia

Dr Martin Ashton-Key,
Medical Advisor, National Commissioning Group, NHS London

Mr John Chapman,
Public contributor

Dr Peter Elton,
Director of Public Health, Bury Primary Care Trust

Dr Ben Goldacre,
Research Fellow, Division of Psychological Medicine and Psychiatry, King's College London

Dr James Gray,
Consultant Microbiologist, Department of Microbiology, Birmingham Children's Hospital NHS Foundation Trust

Ms Kylie Gyertson,
Oncology and Haematology Clinical Trials Manager, Guy's and St Thomas' NHS Foundation Trust London

Dr Jurjees Hasan,
Consultant in Medical Oncology, The Christie, Manchester

Dr Carl Heneghan,
Deputy Director Centre for Evidence-Based Medicine and Clinical Lecturer, Department of Primary Health Care, University of Oxford

Dr Dyfrig Hughes,
Reader in Pharmacoeconomics and Deputy Director, Centre for Economics and Policy in Health, IMSCaR, Bangor University

Dr Maria Kouimtzi,
Pharmacy and Informatics Director, Global Clinical Solutions, Wiley-Blackwell

Professor Femi Oyebode,
Consultant Psychiatrist and Head of Department, University of Birmingham

Dr Andrew Prentice,
Senior Lecturer and Consultant Obstetrician and Gynaecologist, The Rosie Hospital, University of Cambridge

Ms Amanda Roberts,
Public contributor

Dr Martin Shelly,
General Practitioner, Silver Lane Surgery, Leeds

Dr Gillian Shepherd,
Director, Health and Clinical Excellence, Merck Serono Ltd

Mrs Katrina Simister,
Assistant Director New Medicines, National Prescribing Centre, Liverpool

Professor Donald Singer,
Professor of Clinical Pharmacology and Therapeutics, Clinical Sciences Research Institute, CSB, University of Warwick Medical School

Mr David Symes,
Public contributor

Dr Arnold Zermansky,
General Practitioner, Senior Research Fellow, Pharmacy Practice and Medicines Management Group, Leeds University

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Mr Simon Reeve,
Head of Clinical and Cost-Effectiveness, Medicines, Pharmacy and Industry Group, Department of Health

Dr Heike Weber,
Programme Manager, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Psychological and Community Therapies Panel

Members

Chair,
Professor Scott Weich,
Professor of Psychiatry, University of Warwick, Coventry

Deputy Chair,
Dr Howard Ring,
Consultant & University Lecturer in Psychiatry, University of Cambridge

Professor Jane Barlow,
Professor of Public Health in the Early Years, Health Sciences Research Institute, Warwick Medical School

Dr Sabyasachi Bhaumik,
Consultant Psychiatrist, Leicestershire Partnership NHS Trust

Mrs Val Carlill,
Public contributor

Dr Steve Cunningham,
Consultant Respiratory Paediatrician, Lothian Health Board

Dr Anne Hesketh,
Senior Clinical Lecturer in Speech and Language Therapy, University of Manchester

Dr Peter Langdon,
Senior Clinical Lecturer, School of Medicine, Health Policy and Practice, University of East Anglia

Dr Yann Lefeuvre,
GP Partner, Burrage Road Surgery, London

Dr Jeremy J Murphy,
Consultant Physician and Cardiologist, County Durham and Darlington Foundation Trust

Dr Richard Neal,
Clinical Senior Lecturer in General Practice, Cardiff University

Mr John Needham,
Public contributor

Ms Mary Nettle,
Mental Health User Consultant

Professor John Potter,
Professor of Ageing and Stroke Medicine, University of East Anglia

Dr Greta Rait,
Senior Clinical Lecturer and General Practitioner, University College London

Dr Paul Ramchandani,
Senior Research Fellow/Cons. Child Psychiatrist, University of Oxford

Dr Karen Roberts,
Nurse/Consultant, Dunston Hill Hospital, Tyne and Wear

Dr Karim Saad,
Consultant in Old Age Psychiatry, Coventry and Warwickshire Partnership Trust

Dr Lesley Stockton,
Lecturer, School of Health Sciences, University of Liverpool

Dr Simon Wright,
GP Partner, Walkden Medical Centre, Manchester

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Expert Advisory Network

Members

Professor Douglas Altman,
Professor of Statistics in Medicine, Centre for Statistics in Medicine, University of Oxford

Professor John Bond,
Professor of Social Gerontology & Health Services Research, University of Newcastle upon Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery, Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Regulation and Improvement Authority, Belfast

Ms Tracy Bury,
Project Manager, World Confederation of Physical Therapy, London

Professor Iain T Cameron,
Professor of Obstetrics and Gynaecology and Head of the School of Medicine, University of Southampton

Professor Bruce Campbell,
Consultant Vascular & General Surgeon, Royal Devon & Exeter Hospital, Wonford

Dr Christine Clark,
Medical Writer and Consultant Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing and Head of Research, The Medical School, University of Birmingham

Professor Barry Cookson,
Director, Laboratory of Hospital Infection, Public Health Laboratory Service, London

Dr Carl Counsell,
Clinical Senior Lecturer in Neurology, University of Aberdeen

Professor Howard Cuckle,
Professor of Reproductive Epidemiology, Department of Paediatrics, Obstetrics & Gynaecology, University of Leeds

Professor Carol Dezateux,
Professor of Paediatric Epidemiology, Institute of Child Health, London

Mr John Dunning,
Consultant Cardiothoracic Surgeon, Papworth Hospital NHS Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon, Gloucestershire Royal Hospital, Gloucester

Professor Martin Eccles,
Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne

Professor Pam Enderby,
Dean of Faculty of Medicine, Institute of General Practice and Primary Care, University of Sheffield

Professor Gene Feder,
Professor of Primary Care Research & Development, Centre for Health Sciences, Barts and The London School of Medicine and Dentistry

Mr Leonard R Fenwick,
Chief Executive, Freeman Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher,
Antenatal Teacher and Tutor and President, National Childbirth Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine, University of Birmingham

Mr Tam Fry,
Honorary Chairman, Child Growth Foundation, London

Professor Fiona Gilbert,
Consultant Radiologist and NCRN Member, University of Aberdeen

Professor Paul Gregg,
Professor of Orthopaedic Surgical Science, South Tees Hospital NHS Trust

Bec Hanley,
Co-director, TwoCan Associates, West Sussex

Dr Maryann L Hardy,
Senior Lecturer, University of Bradford

Mrs Sharon Hart,
Healthcare Management Consultant, Reading

Professor Robert E Hawkins,
CRC Professor and Director of Medical Oncology, Christie CRC Research Centre, Christie Hospital NHS Trust, Manchester

Professor Richard Hobbs,
Head of Department of Primary Care & General Practice, University of Birmingham

Professor Alan Horwich,
Dean and Section Chairman, The Institute of Cancer Research, London

Professor Allen Hutchinson,
Director of Public Health and Deputy Dean of ScHARR, University of Sheffield

Professor Peter Jones,
Professor of Psychiatry, University of Cambridge, Cambridge

Professor Stan Kaye,
Cancer Research UK Professor of Medical Oncology, Royal Marsden Hospital and Institute of Cancer Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch & Ptnrs), The Health Centre, Thame

Dr Donna Lamping,
Research Degrees Programme Director and Reader in Psychology, Health Services Research Unit, London School of Hygiene and Tropical Medicine, London

Professor James Lindesay,
Professor of Psychiatry for the Elderly, University of Leicester

Professor Julian Little,
Professor of Human Genome Epidemiology, University of Ottawa

Professor Alistaire McGuire,
Professor of Health Economics, London School of Economics

Professor Neill McIntosh,
Edward Clark Professor of Child Life and Health, University of Edinburgh

Professor Rajan Madhok,
Consultant in Public Health, South Manchester Primary Care Trust

Professor Sir Alexander Markham,
Director, Molecular Medicine Unit, St James's University Hospital, Leeds

Dr Peter Moore,
Freelance Science Writer, Ashtead

Dr Andrew Mortimore,
Public Health Director, Southampton City Primary Care Trust

Dr Sue Moss,
Associate Director, Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton

Professor Miranda Mugford,
Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson,
Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick,
Director, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price,
Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg,
Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sandercock,
Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Philip Shackley,
Senior Lecturer in Health Economics, Sheffield Vascular Institute, University of Sheffield

Dr Eamonn Sheridan,
Consultant in Clinical Genetics, St James's University Hospital, Leeds

Dr Margaret Somerville,
Director of Public Health Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown,
Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Dr Nick Summerton,
GP Appraiser and Codirector, Research Network, Yorkshire Clinical Consultant, Primary Care and Public Health, University of Oxford

Professor Ala Szczepura,
Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Dr Ross Taylor,
Senior Lecturer, University of Aberdeen

Dr Richard Tiner,
Medical Director, Medical Department, Association of the British Pharmaceutical Industry

Mrs Joan Webster,
Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle,
Clinical Co-director, National Co-ordinating Centre for Women's and Children's Health, Lymington

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (www.hta.ac.uk) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

NETSCC, Health Technology Assessment
Alpha House
University of Southampton Science Park
Southampton SO16 7NS, UK
Email: hta@hta.ac.uk
www.hta.ac.uk

ISSN 1366-5278