

**Adalimumab, etanercept, infliximab and abatacept for moderate rheumatoid arthritis after conventional DMARDs only have failed (partial review of TA375) [ID2710]**

|  |  |
| --- | --- |
| **Produced by** | The School of Health and Related Research (ScHARR), The University of Sheffield |
| **Authors** | Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK  Emma Simpson, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK  Gill Rooney, Programme Manager, ScHARR, University of Sheffield, Sheffield, UK  Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK  Christopher J Edwards, Consultant Rheumatologist, NIHR Clinical Research Facility, University Hospital Southampton NHS Foundation Trust, Southampton, SO16 6YD |
| **Correspondence Author** | Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK |
| **Date completed** | Date completed (04/01/2021) |

**Source of funding**: This report was commissioned by the NIHR HTA Programme as project number 14/66/01.

**Declared competing interests of the authors**

Chris Edwards declares that he has received honoraria, provided consultancy or been part of a speaker’s bureau for AbbVie, Pfizer, BMS, Celltrion, Fresenius Kabi, Biogen and Janssen in relation to rheumatoid arthritis interventions. None of the remaining authors have any conflicts of interest to declare.

**Acknowledgements**

We would like to thank Andrew Metry, for help in discussing the project.

**Rider on responsibility for report**

'This report was commissioned by the NIHR Evidence Synthesis Programme. The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

**This report should be referenced as follows:**

Stevenson M, Simpson E, Rooney G, Wong R, Edwards CJ. Adalimumab, etanercept, infliximab and abatacept for moderate rheumatoid arthritis after conventional DMARDs only have failed (partial review of TA375) [ID2710]

**Contributions of authors**

Emma Simpson and Gill Rooney undertook the systematic review. Emma Simpson, Gill Rooney and Matt Stevenson evaluated the clinical evidence related to the relationship between changes in HAQ score and DAS28 score. Matt Stevenson amended the health economic model and generated the results. Ruth Wong generated the search strategy used. Chris Edwards provided clinical advice. All authors were involved in drafting and commenting on the final report.

**Standard copyright statement on the front page of the ERG/AG report:**

Copyright belongs to The University of Sheffield.

**Table of Contents**

[1 LIST OF ABBREVIATIONS 1](#_Toc60645239)

[2. EXECUTIVE SUMMARY 3](#_Toc60645240)

[3 BACKGROUND 4](#_Toc60645241)

[3.1 Description of health problem 4](#_Toc60645242)

[3.2 Current service provision 4](#_Toc60645243)

[3.3 Description of technologies under assessment 4](#_Toc60645244)

[4 DEFINITION OF THE DECISION PROBLEM 6](#_Toc60645245)

[5 ASSESSMENT OF CLINICAL EFFECTIVENESS 8](#_Toc60645246)

[5.1 Methods for reviewing effectiveness 8](#_Toc60645247)

[5.2 Results 10](#_Toc60645248)

[6 INDEPENDENT ECONOMIC ASSESSMENT 18](#_Toc60645254)

[6.1 Methods 18](#_Toc60645255)

[6.2 The assumed efficacy of the interventions. 19](#_Toc60645256)

[6.3 Model results 20](#_Toc60645257)

[7 CONCLUSIONS 21](#_Toc60645258)

[8 REFERENCES 22](#_Toc60645259)

[9 APPENDICES 27](#_Toc60645260)

[Appendix 1: Excluded studies with rationale for exclusion 27](#_Toc60645261)

[Appendix 2 Potential sensitivity analyses 28](#_Toc60645262)

**List of Tables**

[Table 1 MEDLINE search strategy: Search conducted October 01 2020 8](#_Toc60645420)

[Table 2 Included study characteristics 12](#_Toc60645421)

[Table 3 Baseline characteristics of included studies 14](#_Toc60645422)

[Table 4 Step changes estimated 16](#_Toc60645423)

[Table 5: Assumed efficacy associated with each treatment, all bDMARDs with MTX 19](#_Toc60645424)

[Table 6 Study characteristics of studies providing data, but excluded 28](#_Toc60645425)

**List of Figures**

[Figure 1 Flow diagram of study selection (based on PRISMA guidelines http://prisma-statement.org/) 11](#_Toc60645515)

[Figure 2: QALYs gained by each strategy 20](#_Toc60645516)

# 1 LIST OF ABBREVIATIONS

|  |  |
| --- | --- |
| ABA | Abatacept |
| ADA | Adalimumab |
| bDMARD | Biologic disease-modifying antirheumatic drug |
| BeST | Behandel Strategieen; in English, treatment strategies |
| CATCH | Canadian Early Arthritis Cohort |
| csDMARD | Conventional Synthetic Disease-Modifying Antirheumatic Drug |
| DAS | Disease Activity Score |
| DAS28 | Disease Activity Score 28 joints |
| DAS28-CRP | Disease Activity Score 28 joints - C-Reactive Protein |
| DAS28-ESR | Disease Activity Score 28 joints - Erythrocyte Sedimentation Rate |
| DAS44 | Disease Activity Score 44 joints |
| DCP | Data from daily clinical practice |
| DMARD | Disease-Modifying Antirheumatic Drug |
| EQ-5D | European Quality of Life 5-Dimensions |
| ERAS | Early Rheumatoid Arthritis Study |
| ETN | Etanercept |
| EULAR | European League Against Rheumatism |
| HAQ | Health Assessment Questionnaire |
| HAQ-DI | Health Assessment Questionnaire Disability Index |
| IFX | Infliximab |
| IQR | Interquartile Range |
| IV | Intravenously |
| J-HAQ | Japanese version of the Health Assessment Questionnaire |
| MTX | Methotrexate |
| NOAR | Norfolk Arthritis Register |
| NSAIDS | Non-Steroidal Anti-Inflammatory Drugs |
| PAS | Patient Access Scheme |
| QALY | Quality-adjusted life years |
| RA | Rheumatoid Arthritis |
| RCT | Randomised Controlled Trial |
| RTX | Rituximab |
| TA | Technology Appraisal |
| T2T | Treat to Target |
| TCZ | Tocilizumab |
| TNFi | Tumour Necrosis Factor inhibitor |
| UK | United Kingdom |
| USA | United States of America |

# 2. EXECUTIVE SUMMARY

This work has been undertaken to partially update NICE technology appraisal 375 (TA375) to consider the cost-effectiveness of biologic disease-modifying antirheumatic drugs (bDMARDs) in patients with moderate-to-severe rheumatoid arthritis (RA). Moderate-to-severe RA is defined as a Disease Activity Score (28 joints) (DAS28) score between 3.2 and 5.1. The manufacturers of four bDMARDs (abatacept, adalimumab, etanercept and infliximab) paid to be considered within the partial update of NICE TA375.

In addition to updating the prices of bDMARDs due to the emergence of biosimilars, the model used for TA375 was updated to account for the fact that patients with moderate-to-severe RA would receive bDMARDs when their RA was deemed severe, with a DAS28 score greater than 5.1. To action this change, the relationship between changes in Health Assessment Questionnaire (HAQ) score and changes in DAS28 scores was required. A systematic search of literature was conducted to source information on this parameter, focussing primarily on people with moderate-to-severe RA. One database was searched: Ovid MEDLINE 1946 to the 1st of October 2020. The systematic review was supplemented by company submissions and papers identified by clinical experts.

Nine studies were identified meeting the inclusion criteria, with data reserved for consideration in sensitivity analyses provided in ten other studies, in subgroups of two of the nine included studies, and from one company submission. Estimates in the change in DAS28 score per 0.125 change in HAQ score was estimated using graphical software where necessary.

There was a wide range in the estimated change in DAS28 score associated with a 0.125 change in HAQ score which ranged from -6.50 to 0.901. The Assessment Group believed that the best estimate was a value of 0.48 which was taken from a study with the intention of estimating the relationship between changes in DAS28 scores and HAQ scores and provided a value near the middle of other estimates. Sensitivity analyses were conducted using a lower value of \*\*\*\*\* to an upper value of 0.90.

Cost-effectiveness results cannot be provided in this document due to the commercial-in-confidence nature of the prices of biosimilars and due to confidential patient access schemes. These results are contained in a confidential addendum.

# 3 BACKGROUND

## 3.1 Description of health problem

Rheumatoid arthritis (RA) is a chronic inflammatory disease which is characterised by progressive and irreversible joint damage, impaired joint function, pain and tenderness caused by swelling of the synovial lining of joints.2 RA is manifested with increasing disability and reduced quality of life.

## 3.2 Current service provision

NICE Technology Appraisal (TA375)3 recommended adalimumab (ADA), etanercept (ETN), infliximab (IFX), certolizumab pegol, golimumab, tocilizumab (TCZ), and abatacept (ABA) in combination with methotrexate (MTX) is recommended for treating patients with RA only if: 1) RA is severe, that is a Disease Activity Score 28 joints (DAS28) score greater than 5.1; 2) the diseases has not responded to intensive therapy with a combination of conventional synthetic disease-modifying antirheumatic drug (csDMARDs); and 3) that the agreed patient access schemes (PAS) for ABA, certolizumab pegol, golimumab and TCZ are provided. ADA, ETN, certolizumab pegol or TCZ can be used as monotherapy for people who cannot take MTX because it is contraindicated or because of intolerance. NICE also stated that treatment should be started with the least expensive drug.

At the time of writing, no biologic disease-modifying antirheumatic drugs (bDMARDs) are recommended by NICE for the treatment of moderate-to-severe RA, which is defined as those patients with a DAS28 score between 3.2 and 5.1. The focus of this partial update is on estimating the cost-effectiveness of bDMARDs for patients with moderate-to-severe RA. Due to the emergence of biosimilars, and the resulting falls in acquisition price for a number of the technologies, it is anticipated that bDMARDs will now be more cost-effective than at the time of TA375.

## 3.3 Description of technologies under assessment

Whilst NICE TA375 provided recommendations on seven interventions, the update only focuses on four: ABA, ADA, ETN, and IFX, as the manufacturers of the omitted interventions did not pay the fee required by NICE for the intervention to be appraised.

ABA is a selective modulator of the T-lymphocyte activation pathway. It binds to molecules on the surface of antigen-presenting cells, preventing full activation of the T lymphocytes and interrupting the inflammatory process. It is provided in two formulations, intravenously (iv) and subcutaneously (sc). The dose regimen for ABA iv is 500 mg below 60 kg, 750 mg between 60 kg and 100 kg, 1000 mg above 100 kg; 0, 2 and 4 weeks, then every 4 weeks thereafter. For ABA sc the dose regimen is 125 mg weekly following a loading dose of 500mg below 60 kg, 750mg between 60 kg and 100 kg, 1000 mg above 100 kg.

ADA, ETN and IFX, all inhibit the activity of tumour necrosis factor alpha, a pro-inflammatory mediator that is partly responsible for damage to the joints in RA. ADA and ETN are provided sc, whereas IFX is an iv administration. ADA is provided at doses of 40mg every other week, ETN at doses of 50mg every week, and IFX is provided at 3mg/kg at weeks 0, 2 and 6 and then every 8 weeks.

All four drugs being appraised are subject to PAS or pricing for biosimilars that are deemed commercial in confidence. As such, the prices cannot be reported in this document, but are contained in a confidential addendum.

# 4 DEFINITION OF THE DECISION PROBLEM

The decision problem is to assess the cost-effectiveness of ABA, ADA, ETN, and IFX when used to treat patients with moderate-to-severe RA compared with the current treatment paradigm. NICE has requested that all parameters values preferred by the NICE Appraisal Committee when it produced guidance for TA375 should be maintained with the exception of two elements which are discussed below.

1. Updating of the prices, where applicable, of ABA, ADA, ETN, IFX, rituximab (RTX) and TCZ.

This has been undertaken to ensure that any price reductions that have occurred since the introduction of biosimilars into the market, or through any changes in PAS are considered. Due to the sequence of interventions modelled, RTX and TCZ are also incorporated as these treatments would be used following discontinuation of the first bDMARD.

1. Amending the mathematical model to ensure that patients with moderate-to-severe RA who do not receive bDMARDs, will receive bDMARDs when their RA becomes severe.

In the model constructed for TA375, patients with moderate-to-severe RA were modelled as having two potential treatment pathways. 1) receive bDMARDs immediately and then progress through a sequence that comprised of RTX, TCZ and then csDMARDs or 2) to forever stay on csDMARDs. This omitted the option for the patient to remain on csDMARDs until their RA became severe, at which point in accordance with NICE recommendations, bDMARDs could be provided. In order to action this change, the model needed to estimate the relationship between changes in Health Assessment Questionnaire (HAQ) score, which was the key metric used in the modelling, and changes in DAS28 score, which is the metric used to determine the severity of RA. The relationship between changes in the parameters were deemed more pertinent for the work than relationships between absolute HAQ and DAS28 scores, as the model explicitly monitors changes in HAQ, which is a scale from zero to 3.0 with steps of 0.125.

Once a relationship between changes in HAQ and changes in DAS28 has been assumed, the amended model monitors the DAS28 score of the patient. If the patient is on the csDMARD-first strategy they will be provided with a bDMARD once the patient reaches a DAS28 score greater than 5.1. Further details of the mechanics of this change are provided in Section 6.

# 5 ASSESSMENT OF CLINICAL EFFECTIVENESS

This chapter details the methods used to identify evidence related to the relationship between changes in HAQ and changes in DAS28, and also presents the results found.

## 5.1 Methods for reviewing effectiveness

A systematic search of literature was conducted to source information on the relationship between the change in HAQ score and the change in DAS28 score.

**Searches**

One database was searched: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) 1946 to the 1st of October, 2020.

The MEDLINE search strategy is shown in Table 1:

**Table 1 MEDLINE search strategy: Search conducted October 01 2020**

|  |  |  |
| --- | --- | --- |
| **#** | **Searches** | **Results** |
| 1 | exp Arthritis, Rheumatoid/ | 113387 |
| 2 | ((rheumatoid or early) adj arthritis).tw. | 107028 |
| 3 | 1 or 2 | 149310 |
| 4 | (("disease activity score" or das\*) adj5 ("health assessment questionnaire" or haq\*)).tw. | 738 |
| 5 | (relationship or associat\* or corrolat\*).tw. | 5263889 |
| 6 | 3 and 4 and 5 | 332 |
| 7 | limit 6 to english language | 328 |

Additionally, references provided within company submissions were checked and papers known to our clinical expert added. The reference lists of relevant studies were checked. All identified citations from the electronic searches and other resources were imported into, and managed using, Endnote X9 software (Clarivate analytics 2020 TM).

**Study selection**

All titles and abstracts were independently examined for inclusion by two reviewers. Any citations that clearly did not meet the inclusion criteria were excluded. Full text articles were sourced and independently checked by two reviewers. Disagreements were resolved by discussion, with involvement of a third member of the team. Study selection was based on the following inclusion and exclusion criteria.

**Inclusion criteria**

***Population***

Adults (aged 18 years and over) with active RA. If data allow, there is a preference for studies reporting on patients with moderate-to-severe RA (DAS28 3.2-5.1). If there are insufficient data, then any severity of RA would be considered.

***Outcome***

Change in HAQ/ Health Assessment Questionnaire Disability Index (HAQ-DI) and the associated change in DAS28 (DAS28-erythrocyte sedimentation rate (DAS28-ESR) or DAS28-c-reactive protein (DAS28-CRP).

***Study design***

Studies were required to provide relevant data, and were not required to be designed solely to address the question of relative changes in HAQ and DAS28.

**Exclusion criteria**

***Population***

Children. Studies of several types of arthritis where data not available separately for RA.

***Outcomes***

Data that cannot be used to calculate change in HAQ and the associated change in DAS28, over the same time period, and in the same group of RA patients. No DAS28 data reported (disease activity score 44 joints (DAS44) is excluded). No HAQ / HAQ-DI data reported (The Japanese version of the Health Assessment Questionnaire (J-HAQ) is excluded as overall disability index higher in the J-HAQ than in the original HAQ4).

***Study design***

Animal models; preclinical and biological studies; narrative reviews, editorials, opinions; and non-English-language papers. Publication type: articles published as abstracts only where insufficient information is available on outcomes or methods.

Where data meeting inclusion criteria are lacking, some allowance may be given (in severity of RA or prior treatment with biologics) for studies to be used in sensitivity analyses.

**Data extraction and synthesis**

Data relevant to the decision problem were extracted by one reviewer, and checked by another. Data were extracted without blinding to authors or journal. Graphical data of change in HAQ or DAS28 were estimated using Engauge software [version 12.1; Mark Mitchell, Los Angeles, CA, USA (2011)]. Data of change in HAQ and DAS28 over the same time period, in the same population of patients, were used to calculate an estimated change in DAS28 for a change in HAQ of 0.125 points.

## 5.2 Results

The MEDLINE search was conducted on the 1st of October 2020. It identified 328 records (without removing duplicates). Twenty-six articles were referenced by company submissions, three articles were recommended by our clinical advisor. The bibliography search yielded four additional articles. The search total, following removal of duplicates was 340 (Figure 1).

Following title/abstract sift, 48 full-text articles were checked. The 29 studies excluded at the full-text stage are listed (with rationale for exclusion) in Appendix 1, leaving 19 studies containing relevant information. Of these nineteen, which nine met the inclusion criteria, and eleven provided data that would be considered for sensitivity analyses if the included studies could not provide sufficient data – two studies provided information for both categories. Results from the studies which met the inclusion criteria are provided in the main text, whereas data for studies which provide data considered for sensitivity analyses are shown in Appendix 2. The reasons for exclusion were having patients with an average baseline DAS28 score > 5.1 (n=8), using DAS44 rather than DAS28, (n=2) and having patients with an average baseline DAS28 score <3.2 (n=1). Additionally, AbbVie provided potentially useful data, although having examined the studies referred to, it appeared probable that these were for patients with an average baseline DAS28 score > 5.1. Data from the AbbVie submission are summarised at a high-level in Appendix 2.

Characteristics of included studies are shown in Table 2. Only one paper (Boyd et al 20135) had a primary outcome to investigate the relationship between function and disease activity over time, and this was a sub-study of the Canadian Early Arthritis Cohort (CATCH). In all nine studies, HAQ and DAS28 were assessed by qualified clinicians (rheumatologists or rheumatology nurses), as part of ongoing patient care, and are unlikely to be subject to biases. As validated, widely used measures, HAQ and DAS28 were not subject to change throughout the follow-up periods of studies.

**Figure 1 Flow diagram of study selection (based on PRISMA guidelines** [**http://prisma-statement.org/**](http://prisma-statement.org/)**)**

Additional records identified through other sources  
(company submissions n=26)

(Bibliography searching n=4)

(Clinical advisor n=3)

Records excluded at title/abstract sift  
(n =292)

Full-text articles assessed for eligibility  
(n=48)

Records screened  
(n = 340)

Records after duplicates removed  
(n=340)

## Identification

## Eligibility

## Included

## Screening

Records identified through database searching  
(n=328)

Full-text articles excluded  
(n =29)

Reasons for exclusion

Data insufficient or unusable (n= 28)

J-HAQ (n=1)

Studies included in narrative synthesis  
(n = 9\*)

Studies with potential for sensitivity analyses

(n=12\*)

\*Two studies provided data for both categories.

Unpublished data provided by company submissions could also be included. This resulted in one additional data set with the potential for use in sensitivity analyses

**Table 2 Included study characteristics**

| **Reference** | **Study type** | **Study objective** | **Sample size** | **Follow-up (months)** |
| --- | --- | --- | --- | --- |
| Ariza-Ariza et al 20066 | Prospective multicentre study | To compare the utility values and quality-adjusted life years (QALYs) obtained by the Time Trade-Off  instrument (TTO) and the European Quality of Life -5 Dimensions (EQ-5D) | 300 | 12 |
| Augustsson et al 20101 | Database study | Investigating Tumour Necrosis Factor inhibitor (TNFi) and workforce participation | 594 | 60 |
| Boyd et al 20135 | Data from Canadian Early Arthritis Cohort (CATCH) | Sub-study investigating function and disease activity in early arthritis | 1,143 | 24 |
| de Andrade et al 20177 | Single centre prospective cohort study | Investigating disease activity and physical function after treat-to-target strategy | 229 | 108 |
| Fioravanti et al 2019 8 | Prospective cohort from two centres in Italy | Investigating TCZ therapy | 44 | 6 |
| Gwinnutt et al 20209 | the Rheumatoid Arthritis Medication Study, a UK multicentre cohort study | Investigating clusters of symptoms associated with poor outcomes in early RA | 1,127 | 12 |
| Ling et al 201610 | data from two cohorts: the Norfolk Arthritis Register (NOAR); and the Early Rheumatoid Arthritis Study (ERAS) | Investigating effect of HLA-DRB1 on disease activity | NOAR n=2,158  ERAS n=329 | 60 |
| Nair et al 201411 | data from clinical practice from the observational Nijmegen Early Rheumatoid Arthritis inception cohort | Investigating whether treatment effects of pragmatic clinical trials are generalisable to data from daily clinical practice (DCP), | DCP n=198 | 6 |
| Twigg et al 201712 | Data from Yorkshire Early Arthritis Register (YEAR) | To assess patient-reported variables as predictors of change in disease activity and disability | 1,415 | 12 |

TNFi=tumour necrosis factor inhibitor; TCZ=tocilizumab; DCP=Data from daily clinical practice; NOAR= Norfolk Arthritis Register; ERAS = Early Rheumatoid Arthritis Study

Baseline variables of included trials are shown in Table 3. Mean/median DAS28 scores were between 3.2 and 5.1 (that is, moderate-to-severe) in all nine studies although, this was only for one of the two cohorts in Nair et al 201411).

Baseline ages were similar across studies, with the lowest mean age 40 years,1 and highest age median 60 years.9 All six studies had a majority of female patients, as is to be expected from prevalence of RA. Baseline disease duration ranged from six months10 5 to 10.6 years.7 This is considered by our clinical advisor to be generalisable to the RA population seeking treatment in England.

The estimated change in DAS28 associated with a 0.125 change in HAQ are provided in Table 4. In all studies apart from Ariza-Ariza *et al*.6 and cluster 6 of Gwinnutt *et al*. 9 HAQ and DAS28 scores decreased indicating an improvement, on average, in the condition of the patients. As such, the assessment group has had to assume that the relationship between decreases in HAQ score and in decreases in DAS28 are generalisable to when there are increases in the HAQ score.

A wide range was observed in the estimated relationship between the change in DAS28 score when HAQ changes. Ariza-Ariza *et al*.6 reported a large, negative correlation whilst a positively correlated estimate of 0.90 was derived from Twigg *et al*.12 The ERG believes that the most appropriate estimate (0.48) would be provided by Boyd *et al*.5 which has the advantage of the relationship being the primary outcome of the study, having a reasonable long follow-up of 24 months, having no bDMARD use, and with an estimate that was not too removed from the remaining studies.

Acknowledging the uncertainty in the parameter the ERG ran two sensitivity analyses using a higher value and a lower value. The higher value (0.90) was estimated from Twigg *et al.*12 which was a fairly recent, large, study of reasonable length without the use of bDMARDs. For the lower value, the ERG preferred to use data reserved for sensitivity analyses and use the values estimated by AbbVie which regressed change in DAS28 on HAQ based on individual patient data from four RCTs of upadacitinib. The reason for choosing this source is that the estimated value (\*\*\*\*\*) is amongst the lowest observed, that individual patient data had been used, and importantly that this was the only source where both HAQ and DAS score was assumed to increase.

**Table 3 Baseline characteristics of included studies**

| **Reference** | **Study Sample size** | **Baseline DAS28\*** | **Baseline HAQ** | **Prior Treatment** | **Treatment during study** | **Baseline age**  **(years)** | **Gender**  **(% female)** | **Baseline disease duration** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Ariza-Ariza *et al.* 20066 | 300 | DAS28-ESR  Mean 4.5  SD1.5 | HAQ  Mean 1.2  SD0.9 | csDMARDs, or bDMARDs at physician discretion | csDMARDs, or bDMARDs at physician discretion | Mean 59.6  SD 13.3 | 82 | Years  Mean 10.3 SD 8.7 |
| Augustsson *et al.* 20101 | 594 | DAS28  Mean 4.7  SD 1.4  N=521 | HAQ  Mean 1.0  SD 0.6  N=528 | No prior bDMARD | First treatment with TNFi  IFX (52.9%)  ETN (34.5%)  ADA (12.6%) | Mean 40.0  SD 9.3 | 66 | Years  Mean 9.4  SD 8.5 |
| Boyd *et al.* 20135 | 1,143 | DAS28  mean 4.53  SD 1.99 | HAQ  Mean 0.94  SD 0.72 | csDMARDs with or without prednisone (physician discretion) or csDMARD naive | csDMARDs with or without prednisone (physician discretion) | Mean 52.2  SD 15.8 | 71.2 | Months  Mean 6.3  SD 3.7 |
| de Andrade *et al.* 20177 | 229 | DAS28  Mean 4.6  SD 1.5 | HAQ-DI  Mean 1.4  SD 0.05 | csDMARD | T2T strategy, two courses of csDMARDs followed by bDMARD (TNFi, with physician discretion for ABA, TCZ, RTX) | Mean 55  SD 11 | 83.8 | Years  Mean 10.6  SD 7.4 |
| Fioravanti *et al.* 20198 | 44 | DAS28-ESR Median 4.630 IQR 4.23-5.25 | HAQ  Median 1.68  IQR 1.04-2.38 | At least two csDMARDs | TCZ (n=20);  TCZ+MTX (n=24) | Median 58.50  IQR 48-69.75 | 86.4 | Years  Median 8  IQR 5-15 |
| Gwinnutt *et al.* 20209 | 1,127 | DAS28-CRP  median 4.1  IQR 3.2, 5.2 | HAQ  Median 1.00  IQR 0.38, 1.63 | MTX naive | Starting MTX | Median 60  IQR 50, 69 | 63.4 | Median 6 months,  IQR 4, 10 |
| Ling *et al.* 201610 | 2,158 NOAR cohort;  329 ERAS cohort | DAS28-ESR  NOAR  Median 3.76  IQR 2.79, 4.78  ERAS  Median 5.06  IQR 4.19, 5.84 | HAQ  NOAR  Median 0.875  IQR 0.25, 1.5  ERAS  Median 1  IQR 0.625, 1.6875 | csDMARDs, and/or corticosteroids | csDMARDs, and/or corticosteroids | Age at symptom onset  NOAR  Median 55  IQR 43–67  ERAS  Median 54  IQR 44–62 | NOAR 65  ERAS 67 | Months  NOAR  Median 6  IQR 3, 12  ERAS  Median 6  IQR 3, 11 |
| Nair *et al.* 201411 | 198 Data from DCP | DAS28  DCP  Mean 5.0  SD 1.3 | HAQ  DCP  Mean 0.8  SD 0.7 | csDMARD naive, no prior corticosteroids | csDMARDs, NSAIDS and/or corticosteroids, and/or biologics | DCP  Mean 54.7  SD 15.2 | DCP 61.3 | <1 year |
| Twigg *et al.* 201712 | 1,415 | DAS28-CRP  Mean 5.01  SD 1.33 | HAQ-DI  Mean 1.22  SD0.57 | csDMARDs | csDMARDs, and/or corticosteroids | Mean 57.7  SD 14.2 | 66 | Months  Mean 7.1  SD 4.3 |

\*unless otherwise stated, unclear if calculated with ESR or CRP

DAS28=Disease Activity Score 28 joints; ESR= erythrocyte sedimentation rate; CRP= C-reactive protein; HAQ = Health Assessment Questionnaire; HAQDI = Health Assessment Questionnaire Disease Index; csDMARD=conventional synthetic disease-modifying antirheumatic drug; NOAR= Norfolk Arthritis Register; ERAS = Early Rheumatoid Arthritis Study; DCP=Data from daily clinical practice;

TNFi=tumour necrosis factor inhibitor; MTX=methotrexate; ABA=abatacept; TCZ=tocilizumab; RTX=rituximab; IQR=interquartile range; IFX=infliximab; ADA=adalimumab; ETN=etanercept; NSAIDS=non-steroidal anti-inflammatories; T2T=treat-to-target

**Table 4 Step changes estimated**

| **Reference** | **Sample size [providing data]** | **Baseline DAS** | **Baseline HAQ** | **Treatment during study** | **Follow-up**  **(Months)** | **Change in DAS associated with 0.125-point change in HAQ** |
| --- | --- | --- | --- | --- | --- | --- |
| Ariza-Ariza *et al.* 20066 | 163 | DAS28-ESR  Mean 4.5  SD1.5 | HAQ  Mean 1.2  SD0.9 | csDMARDs, or bDMARDs at physician discretion | 12 | -6.5  DAS28 decrease  HAQ increase |
| Augustsson *et al.*20101 | 528 | Mean 4.7  SD 1.4  N=521 | Mean 1.0  SD 0.6  N=528 | First treatment with TNFi  IFX (52.9%)  ETN (34.5%)  ADA (12.6%) | 60 | 0.59\*  HAQ and DAS28 score decrease |
| Boyd *et al.* 20135 | 214 | mean 4.53  SD 1.99 | Mean 0.94  SD 0.72 | DMARDs with or without prednisone (physician discretion) | 24 | 0.48\*  HAQ and DAS28 score decrease |
| de Andrade *et al.* 20177 | 229  [156 at year 9] | Mean 4.6  SD 1.5 | Mean 1.4  SD 0.05 | T2T strategy, two courses of csDMARDs followed by biologic (TNFi, with physician discretion for ABA, TCZ, RTX) | 108 | 0.39\*  HAQ and DAS28 score decrease |
| Fioravanti *et al.* 20198 | 44 | Median 4.630 IQR 4.23-5.25 | Median 1.68  IQR 1.04-2.38 | TCZ (n=20);  TCZ+MTX (n=24) | 6 | 0.34  HAQ and DAS28 score decrease |
| Gwinnutt *et al.* 20209 | Cluster 5  71  Cluster 6  46 | DAS28-CRP    Cluster 5  Median 3.4  Cluster 6  Median 3.8  [at month 6 of study – baseline of calculation] | Cluster 5 HAQ  Median 1.5  Cluster 6 HAQ  Median 1.25  [at month 6 of study – baseline of calculation] | Starting MTX | 6  [change from months 6 to 12 of the study] | Cluster 5  0.56  HAQ and DAS28 score decrease  Cluster 6  Not calculable  No change in HAQ, DAS28 score decrease |
| Ling *et al.* 201610 | NOAR 2,158  ERAS 329 | NOAR  Median 3.76  IQR 2.79, 4.78  ERAS  Median 5.06  IQR 4.19, 5.84 | NOAR  Median 0.875  IQR 0.25, 1.5  ERAS  Median 1  IQR 0.625, 1.6875 | csDMARDs, and/or corticosteroids | 60 | NOAR = 0.13  ERAS = 0.11  HAQ and DAS28 score decrease |
| Nair *et al.* 201411 | 198  (DCP) | Mean 5.0  SD 1.3 | HAQ  Mean 0.8  SD 0.7 | csDMARDs, NSAIDS and/or corticosteroids, and/or biologics | 6 | 0.33  HAQ and DAS28 score decrease |
| Twigg *et al.* 201712 | 1,415 | 5.01 | 1.22 | csDMARDs, and/or corticosteroids | 12 | 0.90  HAQ and DAS28 score decrease |

\*estimated from graph

DCP=Data from daily clinical practice; TNFi=tumour necrosis factor inhibitor; NOAR= Norfolk Arthritis Register; ERAS = Early Rheumatoid Arthritis Study;

# 6 INDEPENDENT ECONOMIC ASSESSMENT

## 6.1 Methods

As stated, NICE requested that that all parameters values preferred by the NICE Appraisal Committee when it produced guidance for TA375 should be maintained bar updating the prices of interventions and allowing patients to receive bDMARDs when their DAS28 score was greater than 5.1. Comprehensive details of the modelling approach are provided in Stevenson *et al.*13 In TA375, the sequence after the first bDMARD was accepted, to be the use of RTX and then TCZ, providing TCZ was not used earlier in the treatment sequence then csDMARDs.

The following first line bDMARDs were evaluated: ABA iv, ABA sc, ADA, ETN, and IFX. Each was followed by RTX and then TCZ reverting to csDMARDs following the failure of TCZ. The comparator arm was csDMARDs until a patient reached a DAS score of greater than 5.1 where on the advice of our clinical expert a sequence of ADA, RTX and TCZ was used. For all analyses it was assumed that MTX was used in combination with the bDMARD, and that following TA375 guidance, the results for combination therapy would also be used to generalise to the bDMARDs being used in monotherapy.

The model operationalises the change to bDMARD when the patient has severe RA by calculating the number of HAQ increases, in steps of 0.125, that would be required for the DAS28 score of the patient to be greater than 5.1. Once these net number of HAQ step increases have been reached the patient is assumed to receive ADA.

The model structure has the capacity to run 10 cohorts of patients. Having evaluated early results, the Assessment Group decided that 2 cohorts would be used for the csDMARD strategy, 2 for each of the ADA, ETN, and IFX strategies and 1 each for ABA iv and ABA sc. This was because more precision may be needed for the interventions with biosimilars available as the uncertainty associated with the simulated experience of identical patients (often referred to as first-order uncertainty) would be reduced by apportioning two cohorts.

50,000 patients per cohort were simulated, at that point the Monte Carlo sampling error was low, as for both QALYs and costs, the range between the highest and lowest value from the runs for each bDMARD-first strategy being less than 0.5% of the average value. These variations in costs and QALYs were correlated as younger patients would, on average, accrue both greater QALYs and costs.

Only deterministic results were run as there were shown to be little difference between probabilistic and deterministic results in TA375. Each simulation took in the order of 9 hours to complete.

## 6.2 The assumed efficacy of the interventions.

The assumed efficacy of each intervention used in the model is provided in Table 5. A good European League Against Rheumatism (EULAR) response is better than a moderate EULAR response, which is better than no response. In line with TA375, both ABA sc and RTX was assumed to have the same efficacy of ABA iv.

**Table 5: Assumed efficacy associated with each treatment, all bDMARDs with MTX**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| EULAR response | ABA iv | ADA | csDMARDs | ETA | IFX | TCZ |
| Good | 26.3% | 28.1% | 9.7% | 53.0% | 25.6% | 57.2% |
| Moderate | 41.4% | 40.5% | 35.5% | 32.4% | 42.8% | 33.0% |
| No response | 32.3% | 31.4% | 54.8% | 14.6% | 31.6% | 9.8% |

Further details on the consequences of each EULAR response is provided in Stevenson *et al.*13 If there is no EULAR response to a bDMARD after 6 months the next treatment in the strategy is used.

## 6.3 Model results

As there are biosimilars for RTX, and TCZ has a commercial-in-confidence PAS, full results are not provided here but are supplied to the NICE Appraisal Committee in a confidential addendum. However, to provide some transparency, the QALYs gained by each strategy are provided in Figure 2. For csDMARDs there are three values, each associated with a different relationship between changes in HAQ and change in DAS score. As the results for bDMARDs should not be affected by this parameter, and differences are just Monte-Carlo sampling error, the value for each bDMARD is the average of the three runs. It is seen that when the DAS score of patients with moderate-to-severe RA increases more rapidly, more QALYs are gained due to the earlier use of bDMARDs. However, this would also be associated with additional intervention costs.

**Figure 2: QALYs gained by each strategy**

As stated, the cost-effectiveness results cannot be presented in this document due to commercial-in-confidence pricing. However, the results were not overly sensitive to the choice of parameter value for the relationship between HAQ score changes and DAS28 changes.

# 7 CONCLUSIONS

There appears to be considerable uncertainty in the relationship between changes in HAQ scores and changes in DAS28 scores. A limitation within the published literature is that HAQ score was increasing in only one study; as such the Assessment Group had to assume that the relationship associated with decreasing HAQ scores would also apply when HAQ scores increased.

Our best estimate (0.48) is that reported by Boyd *et al.*5 which was a study designed for the purpose of establishing such a relationship and provided a value near the middle of other estimates. Sensitivity values were provided for higher (0.90) and lower values (\*\*\*\*\*) for this relationship.

Cost-effectiveness results cannot be provided in this document, but the incremental cost-effectiveness results were not overly sensitive to the assumed relationship between change in HAQ score and change in DAS28 score.

# 8 REFERENCES

1. Augustsson J, Neovius M, Cullinane-Carli C, Eksborg S, van Vollenhoven RF. Patients with rheumatoid arthritis treated with tumour necrosis factor antagonists increase their participation in the workforce: potential for significant long-term indirect cost gains (data from a population-based registry). *Annals of the Rheumatic Diseases* 2010;69:126-31.

2. Scott DL, Steer S. The course of established rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2007;21:943-67.

3. National Institiute for Health and Care Excellence. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed. 2016. <https://www.nice.org.uk/guidance/ta375> (Accessed 23/12/2020).

4. Matsuda Y, Singh G, Yamanaka H, Tanaka E, Urano W, Taniguchi A*, et al.* Validation of a Japanese version of the Stanford Health Assessment Questionnaire in 3,763 patients with rheumatoid arthritis. *Arthritis Rheum* 2003;49:784-8.

5. Boyd TA, Bonner A, Thorne C, Boire G, Hitchon C, Haraoui BP*, et al.* The Relationship Between Function and Disease Activity as Measured by the HAQ and DAS28 Varies Over Time and by Rheumatoid Factor Status in Early Inflammatory Arthritis (EIA). Results from the CATCH Cohort. *The open rheumatology journal* 2013;7:58-63.

6. Ariza-Ariza R, Hernandez-Cruz B, Carmona L, Dolores Ruiz-Montesinos M, Ballina J, Navarro-Sarabia F*, et al.* Assessing utility values in rheumatoid arthritis: a comparison between time trade-off and the EuroQol. *Arthritis & Rheumatism* 2006;55:751-6.

7. de Andrade NPB, da Silva Chakr RM, Xavier RM, Viecceli D, Correa RHB, de Oliveira Filho CM*, et al.* Long-term outcomes of treat-to-target strategy in established rheumatoid arthritis: a daily practice prospective cohort study. *Rheumatology International* 2017;37:993-7.

8. Fioravanti A, Tenti S, Bacarelli MR, Damiani A, Li Gobbi F, Bandinelli F*, et al.* Tocilizumab modulates serum levels of adiponectin and chemerin in patients with rheumatoid arthritis: potential cardiovascular protective role of IL-6 inhibition. *Clinical & Experimental Rheumatology* 2019;37:293-300.

9. Gwinnutt JM, Hyrich KL, Lunt M, Co-Investigators R, Barton A, Verstappen SMM. Long-term outcomes of patients who rate symptoms of rheumatoid arthritis as 'satisfactory'. *Rheumatology* 2020;59:1853-61.

10. Ling SF, Viatte S, Lunt M, Van Sijl AM, Silva-Fernandez L, Symmons DP*, et al.* HLA-DRB1 Amino Acid Positions 11/13, 71, and 74 Are Associated With Inflammation Level, Disease Activity, and the Health Assessment Questionnaire Score in Patients With Inflammatory Polyarthritis. *Arthritis & Rheumatology* 2016;68:2618-28.

11. Nair SC, Kievit W, Janse RW, Bijlsma JW, Fransen J, Lafeber FP*, et al.* Generalization and extrapolation of treatment effects from clinical studies in rheumatoid arthritis. *Arthritis care & research* 2014;66:998-1007.

12. Twigg S, Hensor EMA, Emery P, Tennant A, Morgan AW, Yorkshire Early Arthritis Register C. Patient-reported Outcomes as Predictors of Change in Disease Activity and Disability in Early Rheumatoid Arthritis: Results from the Yorkshire Early Arthritis Register. *Journal of Rheumatology* 2017;44:1331-40.

13. Stevenson M, Archer R, Tosh J, Simpson E, Everson-Hock E, Stevens J*, et al.* Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. *Health Technol Assess* 2016;20:1-610.

14. Adams R, Walsh C, Veale D, Bresnihan B, FitzGerald O, Barry M. Understanding the relationship between the EQ-5D, SF-6D, HAQ and disease activity in inflammatory arthritis. *Pharmacoeconomics* 2010;28:477-87.

15. Baganz L, Richter A, Albrecht K, Schneider M, Burmester GR, Zink A*, et al.* Are prognostic factors adequately selected to guide treatment decisions in patients with rheumatoid arthritis? A collaborative analysis from three observational cohorts. *Seminars in Arthritis & Rheumatism* 2019;48:976-82.

16. Bechman K, Tweehuysen L, Garrood T, Scott DL, Cope AP, Galloway JB*, et al.* Flares in Rheumatoid Arthritis Patients with Low Disease Activity: Predictability and Association with Worse Clinical Outcomes. *Journal of Rheumatology* 2018;45:1515-21.

17. Bergstra SA, Couto MC, Govind N, Chopra A, Salomon Escoto K, Murphy E*, et al.* Impact of the combined presence of erosions and ACPA on rheumatoid arthritis disease activity over time: results from the METEOR registry. *RMD Open* 2019;5:e000969.

18. Boers M, Buttgereit F, Saag K, Alten R, Grahn A, Storey D*, et al.* What Is the Relationship Between Morning Symptoms and Measures of Disease Activity in Patients With Rheumatoid Arthritis? *Arthritis care & research* 2015;67:1202-9.

19. Bremander A, Forslind K, Eberhardt K, Andersson MLE. Importance of Measuring Hand and Foot Function Over the Disease Course in Rheumatoid Arthritis: An Eight-Year Follow-Up Study. *Arthritis care & research* 2019;71:166-72.

20. Campbell RC, Batley M, Hammond A, Ibrahim F, Kingsley G, Scott DL. The impact of disease activity, pain, disability and treatments on fatigue in established rheumatoid arthritis. *Clinical Rheumatology* 2012;31:717-22.

21. Carvalho PD, Ferreira RJO, Landewe R, Vega-Morales D, Salomon-Escoto K, Veale DJ*, et al.* Association of 17 Definitions of Remission with Functional Status in a Large Clinical Practice Cohort of Patients with Rheumatoid Arthritis. *Journal of Rheumatology* 2020;47:20-7.

22. Drossaers-Bakker KW, de Buck M, van Zeben D, Zwinderman AH, Breedveld FC, Hazes JM. Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. *Arthritis Rheum* 1999;42:1854-60.

23. Fatima S, Schieir O, Valois MF, Bartlett SJ, Bessette L, Boire G*, et al.* Health Assessment Questionnaire at One Year Predicts All-Cause Mortality in Patients with Early Rheumatoid Arthritis. *Arthritis & Rheumatology* 2020;06:06.

24. Heinimann K, von Kempis J, Sauter R, Schiff M, Sokka-Isler T, Schulze-Koops H*, et al.* Long-Term Increase of Radiographic Damage and Disability in Patients with RA in Relation to Disease Duration in the Era of Biologics. Results from the SCQM Cohort. *Journal of Clinical Medicine* 2018;7:13.

25. Lee KE, Choi SE, Xu H, Kang JH, Park DJ, Lee SS. HAQ score is an independent predictor of sustained remission in patients with rheumatoid arthritis. *Rheumatology International* 2017;37:2027-34.

26. Linde L, Sorensen J, Ostergaard M, Horslev-Petersen K, Rasmussen C, Jensen DV*, et al.* What factors influence the health status of patients with rheumatoid arthritis measured by the SF-12v2 Health Survey and the Health Assessment Questionnaire? *Journal of Rheumatology* 2009;36:2183-9.

27. Michaud K, Wallenstein G, Wolfe F. Treatment and nontreatment predictors of health assessment questionnaire disability progression in rheumatoid arthritis: a longitudinal study of 18,485 patients. *Arthritis care & research* 2011;63:366-72.

28. Nikiphorou E, Norton S, Young A, Carpenter L, Dixey J, Walsh DA*, et al.* Association between rheumatoid arthritis disease activity, progression of functional limitation and long-term risk of orthopaedic surgery: combined analysis of two prospective cohorts supports EULAR treat to target DAS thresholds. *Annals of the Rheumatic Diseases* 2016;75:2080-6.

29. Norton S, Sacker A, Dixey J, Done J, Williams P, Young A*, et al.* Trajectories of functional limitation in early rheumatoid arthritis and their association with mortality. *Rheumatology* 2013;52:2016-24.

30. Norton S, Fu B, Scott DL, Deighton C, Symmons DPM, Wailoo AJ*, et al.* Health Assessment Questionnaire disability progression in early rheumatoid arthritis: systematic review and analysis of two inception cohorts. *Seminars in arthritis and rheumatism* 2014;44:131-44.

31. Pan Y, Norton S, Gwinnutt JM, Kearsley-Fleet L, Symmons DPM, Lunt M*, et al.* Not all moderate disease is the same - Identification of disability trajectories among patients with rheumatoid arthritis and moderate disease activity. *PLoS ONE [Electronic Resource]* 2019;14:e0215999.

32. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44-8.

33. Rydell E, Forslind K, Nilsson JA, Jacobsson LTH, Turesson C. Smoking, body mass index, disease activity, and the risk of rapid radiographic progression in patients with early rheumatoid arthritis. *Arthritis Research & Therapy* 2018;20:82.

34. Scott DL, Pugner K, Kaarela K, Doyle DV, Woolf A, Holmes J*, et al.* The links between joint damage and disability in rheumatoid arthritis. *Rheumatology (Oxford)* 2000;39:122-32.

35. Shadick NA, Gerlanc NM, Frits ML, Stolshek BS, Brady BL, Iannaccone C*, et al.* The longitudinal effect of biologic use on patient outcomes (disease activity, function, and disease severity) within a rheumatoid arthritis registry. *Clinical Rheumatology* 2019;38:3081-92.

36. Sokka T, Kankainen A, Hannonen P. Scores for functional disability in patients with rheumatoid arthritis are correlated at higher levels with pain scores than with radiographic scores. *Arthritis Rheum* 2000;43:386-9.

37. Tanaka E, Mannalithara A, Inoue E, Hara M, Tomatsu T, Kamatani N*, et al.* Efficient management of rheumatoid arthritis significantly reduces long-term functional disability. *Annals of the Rheumatic Diseases* 2008;67:1153-8.

38. Ten Klooster PM, Oude Voshaar MAH, Fakhouri W, de la Torre I, Nicolay C, van de Laar M. Long-term clinical, functional, and cost outcomes for early rheumatoid arthritis patients who did or did not achieve early remission in a real-world treat-to-target strategy. *Clinical Rheumatology* 2019;38:2727-36.

39. van der Heijde D, Klareskog L, Singh A, Tornero J, Melo-Gomes J, Codreanu C*, et al.* Patient reported outcomes in a trial of combination therapy with etanercept and methotrexate for rheumatoid arthritis: the TEMPO trial. *Ann Rheum Dis* 2006;65:328-34.

40. Ward MM, Guthrie LC, Alba MI. Measures of arthritis activity associated with patient-reported improvement in rheumatoid arthritis when assessed prospectively versus retrospectively. *Arthritis care & research* 2015;67:776-81.

41. Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis & Rheumatism* 2001;44:2009-17.

42. Tanaka E, Mannalithara A, Inoue E, Iikuni N, Taniguchi A, Momohara S*, et al.* Effects of long-term corticosteroid usage on functional disability in patients with early rheumatoid arthritis, regardless of controlled disease activity. *Rheumatology International* 2012;32:749-57.

43. Genovese MC, Kremer J, Zamani O, Ludivico C, Krogulec M, Xie L*, et al.* Baricitinib in Patients with Refractory Rheumatoid Arthritis. *New England Journal of Medicine* 2016;374:1243-52.

44. Wendler J, Burmester GR, Sorensen H, Krause A, Richter C, Tony HP*, et al.* Rituximab in patients with rheumatoid arthritis in routine practice (GERINIS): six-year results from a prospective, multicentre, non-interventional study in 2,484 patients. *Arthritis Research & Therapy* 2014;16:R80.

45. Koizumi R, Koyama K, Wako M, Ohba T, Takayama Y, Haro H. Clinical conditions needed to acquire sustained functional remission in rheumatoid arthritis patients. *Clinical Rheumatology* 2020; 10.1007/s10067-020-05451-3.

46. Baker JF, Conaghan PG, Emery P, Baker DG, Ostergaard M. Relationship of patient-reported outcomes with MRI measures in rheumatoid arthritis. *Annals of the Rheumatic Diseases* 2017;76:486.

47. van der Kooi E, Klarenbeek NB, Guler-Yuksel M, Kerstens PJ, van der Lubbe PA, Westedt ML*, et al.* A decrease in disease activity score (DAS) level is associated with a decrease in health assessment questionnaire (HAQ) score, independent of follow-up duration, during 5 years of tightly controlled treatment: results from the BeSt study. *Annals of the Rheumatic Diseases* 2011;70:168-71.

48. Andersson MLE, Forslind K, Hafstrom I, Group BS. Patients with Early Rheumatoid Arthritis in the 2000s Have Equal Disability and Pain Despite Less Disease Activity Compared with the 1990s: Data from the BARFOT Study over 8 Years. *Journal of Rheumatology* 2017;44:723-31.

49. Behrens F, Koehm M, Schwaneck EC, Schmalzing M, Gnann H, Greger G*, et al.* Use of a “critical difference” statistical criterion improves the predictive utility of the Health Assessment Questionnaire-Disability Index score in patients with rheumatoid arthritis. *BMC Rheumatology* 2019;3:51.

50. Nair SC, Bijlsma JW, van der Werf JH, van der Veen MJ, Linn-Rasker SP, Vreugdenhil S*, et al.* Do radiographic joint damage and disease activity influence functional disability through different mechanisms? Direct and indirect effects of disease activity in established rheumatoid arthritis. *Journal of Rheumatology* 2013;40:1505-12.

51. Radner H, Alasti F, Smolen JS, Aletaha D. Physical function continues to improve when clinical remission is sustained in rheumatoid arthritis patients. *Arthritis Research & Therapy* 2015;17:203.

# 9 APPENDICES

## Appendix 1: Excluded studies with rationale for exclusion

Full-text articles excluded (n =29)

**Reasons for exclusion**

***Data insufficient or unusable*** *(n= 28)*

Adams *et al*. 201014

Baganz *et al*. 201915

Bechman *et al*. 201816

Bergstra *et al*. 201917

Boers *et al*. 201518

Bremander *et al*. 201919

Campbell *et al*. 201220

Carvalho *et al*. 202021

Drossaers-Bakker *et al*. 199922

Fatima *et al*. 202023

Heinimann *et al*. 201824

Lee *et al*. 201725

Linde *et al*. 200926

Michaud *et al*. 201127

Nikiophorou *et al*. 201628

Norton *et al*. 201329

Norton *et al*. 201430

Pan *et al*. 201931

Prevoo *et al*. 199532

Rydell *et al*. 201833

Scott *et al*. 200034

Shadick *et al*. 201935

Sokka *et al*. 200036

Tanaka *et al*. 200837

Ten Klooster *et al*. 201938

van der Heijde *et al*. 200639

Ward *et al*. 201540

Welsing *et al*. 200141

***J-HAQ*** *(n=1)*

Tanaka *et al*. 201242

## Appendix 2 Potential sensitivity analyses

Studies with potential for inclusion in sensitivity analyses

Eleven published studies were considered for sensitivity analyses if the included studies could not provide sufficient data. Data from a cohort presented in one of the included studies, Nair et al 2014 11 had baseline DAS over 5.1 and is presented in this appendix. Details are provided in Table 6.

Most of the studies had a population with no prior biologic treatment, however three studies included patients with prior biologics at baseline (Genovese et al 201643 Wendler et al 2014.44 Koizumi et al 202045).

In all eleven studies, HAQ and DAS were assessed by physicians. Blinding of outcome assessors was explicit in two studies46 43 and a third study had DAS calculations by a blinded research nurse.47

Furthermore, unpublished data presented by AbbVie in its submission to NICE has been included in Table 6.

**Table 6 Study characteristics of studies providing data, but excluded**

| **Reference** | **Study** | **Sample size** | **Follow-up** | **Reason not meeting inclusion criteria** |
| --- | --- | --- | --- | --- |
| Abbvie unpublished data | “*upadacitinib trials’ data*” | >1000 | ≥ 3 months | Baseline DAS28 score > 5.1 |
| Andersson *et al*. 201748 | Comparing outcomes of two cohorts of RA patients, data from the BARFOT study | Cohort 1 n=928  Cohort 2 n=1010 | 8 years | Baseline DAS28 score > 5.1 |
| Baker *et al*. 201746 | MRI sub-study of GOBEFORE, RCT of golimumab among methotrexate-naïve patients | 291 | 12 months | Baseline DAS28 score > 5.1 |
| Behrens *et al*. 201949 | Data from multicentre observational trial, full cohort and restricted cohort data, to determine a statistically  defined critical difference for HAQ-DI | 2740 | 6 months | Baseline DAS28 score > 5.1 (for cohort providing data) |
| Genovese *et al*. 201643 | Investigating baricitinib treatment, RA-BEACON RCT | 527 | 24 weeks | Baseline DAS28 score > 5.1 |
| Gwinnutt *et al*. 20209  Clusters 1-4 | Investigating clusters of symptoms associated with poor outcomes in early RA in the Rheumatoid Arthritis Medication Study, a UK multicentre cohort study | 455 | 12 months 6 months of HAQ and DAS28 score changes | Baseline DAS28 score > 5.1 |
| Koizumi *et al*. 202045 | Investigating factors for maintaining long-term functional remission, data from database of patient records | 205  (of whom Remission n=154;  No remission n=51) | 1 year | Baseline DAS28 score < 3.2.  Data from Japanese treatment, and so probably J-HAQ (not HAQ) |
| Nair *et al*. 201350 | Investigating disease activity and  functional disability in T2T of RA, data from three cohorts, Netherlands | 1, 034  (of whom Pyramid cohort n=551; CAMERA I n=299; CAMERA II n=236) | 120 months | Baseline DAS28 score > 5.1 |
| Nair *et al*. 201411 | Investigating whether treatment effects of pragmatic clinical trials are generalisable to clinical practice,  data from pragmatic clinical trials of the Utrecht Rheumatoid Arthritis Cohort | Data from RCTs n=398; | 6 months | Baseline DAS28 score > 5.1 |
| Norton *et al*. 201329 | to identify subgroups with distinct trajectories of functional (HAQ) progression, Consecutive patients diagnosed with RA with symptoms <2 years (median 6 months) and prior to disease-modifying treatment were recruited into the Early RA Study (ERAS) | 1460 | 10 years | DAS44 used (not DAS28) |
| Radner *et al*. 201551 | investigating the  course of physical function in patients with sustained (24 weeks) DAS28 remission (DAS28CRP≤2.6),  Information from clinical trials in RA patients and newly introduced TNFi or csDMARDs | 610 | 24 weeks | Baseline DAS28 score > 5.1 |
| van der Kooi *et al*. 201147 | Investigating DAS and functional ability during DAS-steered treatment, data from BeST RCT | 508 | 5years | DAS44 used (not DAS28) |
| Wendler *et al*. 201444 | Investigating RTX in RA, prospective observational study (GERINIS study) | 1658 | 8 months | Baseline DAS28 score > 5.1 |