Therapeutic monitoring of TNF-alpha inhibitors in rheumatoid arthritis

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Irina Tikhonova, *Senior Research Fellow*[1]
Huiqin Yang, *Senior Research Fellow*[1]
Andy Salmon, *Associate Research Fellow*[2]
Segun Bello, *Postdoctoral Research Associate*[1]
Jaime Peters, *Senior Research Fellow*[1]
Sophie Robinson, *Information Specialist*[1]
Mohsen Rezaei, *Research Fellow*[1]
Sophie Dodman, *Research Assistant*[1]
Helen Coelho, *Research Fellow*[1]
Andriy Kharechko, *Postdoctoral Research Associate*[1]
Rich Haigh, *Consultant Rheumatologist*[3]
Meghna Jani, *NIHR Clinical Lecturer in Rheumatology*[4]
Timothy J McDonald, *Consultant Clinical Scientist and Clinical Associate Professor*[3,5]
Martin Pitt, *Associate Professor*[2]
Martin Hoyle, *Associate Professor of Health Technology Assessment*[1]
Stuart Logan, *Cerebra Professor of Paediatric Epidemiology*[5]

[1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School
[2] PenCHORD, University of Exeter Medical School
[3] Royal Devon and Exeter hospital
[4] University of Manchester
[5] University of Exeter Medical School

Correspondence to

Dr Irina Tikhonova
Peninsula Technology Assessment Group (PenTAG)

University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU
[redacted]

Date completed 07/01/2019

Source of funding This report was commissioned by the NIHR HTA Programme as project number 17/10/02.

Declared competing interests of the authors Dr Meghna Jani declared receipt of speaker fees from Grifols. The remaining authors listed declared no competing interests.

Acknowledgments We would like to thank Rob Anderson for his support in preparation of the protocol and in the finalisation of the report, Ruben Mujica Mota for providing comments on the report, Jo Thompson-Coon for overall support of the project, and Louise Crathorne for support during the finalisation of the project (all from the University of Exeter Medical School). We are grateful to Sean Gavan and Katherine Payne from the University of Manchester for their advice on economic assessment of TNF testing; Grant Smith, Nicola Finch, Leanne Brown, Keith Oldfield and Rob Storey from the Royal Devon & Exeter NHS Foundation Trust for their advice on costing. We acknowledge the excellent administrative support of Sue Whiffin and Jenny Lowe.

This report should be referenced as follows: Tikhonova I, Yang H, Salmon A, Bello S, Peters J, Robinson S, Rezaei M, Dodman S, Coelho H, Kharechko A, Haigh R, Jani M, McDonald TJ, Pitt M, Hoyle M, Logan S. Therapeutic monitoring of TNF-alpha inhibitors in rheumatoid arthritis: 2019, Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School (Report for NICE).

Contributions of authors

Segun Bello	Assisted with tasks on development of the protocol for the *Assessment of clinical effectiveness* section, screened records for the clinical-effectiveness systematic review, conducted data extraction for included studies, conducted study quality assessment, contributed to drafting and editing of the *Assessment of clinical effectiveness* section
Helen Coelho	Wrote the *Background* section of the report
Sophie Dodman	Conducted a review of literature to inform costs, contributed to the writing of the *Resources and costs* section
Rich Haigh	Provided clinical advice and contributed to the editing of the report
Martin Hoyle	Contributed to the development of the NICE Scope and Protocol and was the Project Director till 29 June, 2018
Meghna Jani	Provided clinical advice, academic input regarding therapeutic drug monitoring and contributed to the editing of the report
Andriy Kharechko	Conducted a review of literature to inform utilities, contributed to the writing of the *Utilities* section
Stuart Logan	Director and Guarantor of the report
Timothy McDonald	Provided clinical advice and contributed to the editing of the report
Jaime Peters	Contributed to the development of the PenTAG independent economic assessment, including screening titles and abstracts for an additional systematic review supporting the PenTAG independent economic assessment; contributed to the writing and editing of the report
Martin Pitt	Contributed to the development of the PenTAG independent economic assessment
Mohsen Rezaei	Contributed to the cost-effectiveness systematic review as the first reviewer, drafted the *Systematic review of existing cost-effectiveness evidence* and *Independent economic assessment* sections, contributed to the writing of the model code
Sophie Robinson	Developed and conducted the literature searches for the systematic reviews of clinical effectiveness and cost effectiveness, and for the additional systematic review for comparator tests; developed and conducted the searches for the literature review of health utilities; contributed to the writing, editing and proofreading of the report
Andrew Salmon	Contributed to the cost-effectiveness systematic review as the second reviewer, screening of titles and abstracts for the additional systematic review supporting the PenTAG independent economic assessment; contributed to the writing of the *Systematic review of existing cost-effectiveness evidence* and *Independent economic assessment* sections; checked the working of the simulation model
Irina Tikhonova	Provided overall project management; led the health-economic and modelling team; contributed to the writing of the *Abstract, Scientific summary*, *Systematic review of existing cost-effectiveness evidence*, *Independent economic assessment, Discussion* and *Conclusion* sections, and the model code; collated the report
Huiqin Yang	Led the clinical effectiveness of the project, and the writing of the *Abstract* and *Assessment of clinical effectiveness* sections; contributed to the writing of the *Scientific summary, Discussion* and *Conclusion* sections of the report

Please refer to the International Committee for Medical Journal Editors (ICMJE) Uniform Requirements for Manuscripts submitted to Biomedical Journals see http://www.icmje.org/

Abstract

Background

Rheumatoid arthritis (RA) is a chronic systemic autoimmune disease, primarily causing inflammation, pain and stiffness (synovitis) in the joints. Those with severe disease may be treated with biological disease modifying anti-rheumatic drugs (bDMARDs), including TNF-α inhibitors. Monitoring response to these treatments typically involves clinical assessment and the use of response criteria (DAS28 or EULAR).

Commercial enzyme-linked immunosorbent assay (ELISA) tests can also now be used to detect and measure drug concentrations and drug antibody levels in the blood. These tests may inform whether adjustments to treatment are required, or help clinicians to understand the reasons for treatment non-response or a loss of response.

Methods

A systematic review was conducted to identify studies reporting the clinical and cost-effectiveness of using ELISA tests to measure drug levels and anti-drug antibodies for monitoring response to TNF-α inhibitors (adalimumab [ADL], etanercept [ETN], infliximab [IFX], certolizumab pegol [CTZ], and golimumab [GLM]) in people with RA who had achieved treatment target (remission or low disease activity [LDA]), or in those with primary non-response or a secondary non-response to treatment.

An economic analysis was conducted to estimate health and economic outcomes of adding TNF-α inhibitors testing to usual practice to guide treatment decisions. The costs and resource use were considered from the perspective of the NHS and Personal Social Services. No discounting was applied to costs or effects due to the short-term time horizon used. Sensitivity analyses explored the effect of different uncertainties on the economic outcomes.

Results

Eight studies (in 11 publications) were identified. One non-randomised trial (the INGEBIO study, only reported in three abstracts) compared TDM with standard care had serious limitations in relation to the NICE scope: one-third of participants with RA, analyses were mostly not by intention-to-treat, follow-up only 18 months, and, there was no explicit algorithm for guiding clinicians in using the test results to inform treatment. Also, seven observational studies (eight publications) were identified but were of minimal value in informing whether ELISA test-based monitoring is clinically effective or not.

The exploratory economic analyses were inconclusive and show considerable uncertainty in the cost-effectiveness of TDM of TNF-alpha inhibitors in RA. Different outcome data from the same study produced opposite conclusions on the cost-effectiveness of Promonitor testing in people receiving ADL who are in remission/LDA. Results based on the longer follow-up outcomes suggested that monitoring is more costly and produces fewer QALYs than standard care. Of the sensitivity analyses conducted, only the impact of monitoring on the rate of flares impacted substantially on the results. Exploratory analyses of using Promonitor to monitor patients in remission/LDA receiving ETN or INF showed the same results as that for ADL.

Conclusions

There is limited valid and applicable research evidence, and much uncertainty in relation to key potential drivers of the effectiveness and cost-effectiveness of using ELISA-based testing to monitor treatment with bDMARDs in people with RA. Therefore no firm conclusions are possible.

Funding

This report was commissioned by the NIHR HTA Programme as project number 17/10/02.

Word count: 492 words

Scientific summary

Background

Rheumatoid arthritis (RA) is a chronic systemic autoimmune disease, primarily causing inflammation, pain and stiffness (synovitis) in the joints. The course of RA varies considerably from person to person, but often results in substantial morbidity, impaired physical activity, poor quality of life, and reduced life-expectancy. Those with severe disease may be treated with biological disease modifying anti-rheumatic drugs (bDMARDs), including the tumour necrosis factor-alpha (TNF-α) inhibitors adalimumab (ADL), etanercept (ETN), infliximab (IFX), certolizumab pegol (CTZ) and golimumab (GLM). These treatments are costly (from about £5,750 to £9,330 for originator drugs per patient per year) and are commonly associated with side effects including increased risk of infections, allergic reactions, nausea and vomiting, itching, and fever.

Monitoring response to treatment may enable treatment adjustments to be made, thus preventing unnecessary treatment and potentially reducing costs and side effects. Monitoring of response to treatment with TNF-α inhibitors typically involves clinical assessment and the use of the disease activity score in 28 joints (DAS28) or European League Against Rheumatism (EULAR) response criteria. More recently, biochemical ELISA testing has emerged to monitor blood levels of TNF-α inhibitors, or antibodies to TNF-α inhibitors in people with RA.

Objectives

This assessment aims to evaluate the clinical effectiveness and cost-effectiveness of using ELISA tests (Promonitor, IDKmonitor, LISA-TRACKER, RIDASCREEN, MabTrack, and ELISA tests used by Sanquin Diagnostic Services) to measure drug levels and anti-drug antibodies for monitoring response to TNF-α inhibitors (ADL, ETN. IFX, CTZ, GLM) in people with RA who had achieved treatment target (remission or low disease activity [LDA]), or in those who have experienced a primary non-response or a secondary non-response to treatment.

Methods

Assessment of clinical effectiveness

The systematic review was conducted following CRD and NICE guidelines on the conduct of systematic reviews. We performed the systematic review according to a pre-specified

protocol which was registered on the international prospective register of systematic reviews (PROSPERO: CRD42018105195).

The following bibliographic databases were searched: MEDLINE, MEDLINE In-Process, EMBASE, Cochrane Central Register of Controlled Trials (CENTRAL), CDSR, DARE, Web of Science, Clinical Trials.gov, WHO Registry and EU trials from inception to July 2018, and again in November 2018. These searches were supplemented by consultation with experts in the field, and reference-checking of relevant systematic reviews and included studies.

Inclusion criteria

Population: For the systematic review of clinical effectiveness, the eligible populations were people with RA who were being treated with a TNF-α inhibitor (ADL, ETN, IFX, CTZ or GLM) and had achieved treatment target (remission or LDA), or had experienced a primary non-response or a secondary non-response.

Interventions and comparators: The following eligible enzyme linked immunosorbent assay (ELISA) test kits or diagnostic services for monitoring response to TNF-α inhibitors were included: Promonitor ELISA kits, IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits, MabTrack ELISA kits and Sanquin Diagnostic Services. These intervention tests were used in addition to current clinical practice (clinical assessment and monitoring using a composite score such as the DAS28). The comparator considered was standard monitor for people with RA bring treated with a TNF-α inhibitor where treatment decisions were based on clinical judgements and monitoring using a composite score such as DAS28 without the knowledge of circulating drug levels and anti-drug antibodies by means of ELISA tests.

Outcomes: The eligible patient-related outcomes included change in disease activity, change in disease response, rates of hospitalisation, rates of surgical intervention, adverse events (AEs), and health-related quality of life (HRQoL). The clinically important intermediate outcomes included change in number, direction and magnitude of anti-TNF dose, discontinuation of ineffective therapy, and change in frequency of dose adjustment due to monitoring response.

Study designs: Both randomised controlled trials (RCTs) and non-randomised controlled studies comparing therapeutic drug monitoring (TDM) by using ELISA tests with standard care were sought.

Two researchers independently screened the titles and abstracts of all reports identified by the search strategy and full-text papers were subsequently obtained for assessment. Data

extraction and quality assessment were undertaken by one researcher and checked by a second. If RCTs had been identified, the Cochrane Risk of Bias tool for RCTs would have been used. The risk of bias was assessed using the Cochrane (ROBINS-1) tool for non-randomised studies with adaptations as appropriate.

For clinical effectiveness outcomes, mean differences, relative risks (RRs), odds ratios (ORs) or incidence rate ratios (IRRs) with 95% confidence intervals (CIs) were extracted from comparative studies, where reported. Quantitative methods of synthesis were not possible given the considerable clinical heterogeneity associated with interventions, outcomes and length of follow-up and the methodological heterogeneity identified (e.g. different study designs), data were synthesised in narratively (i.e. using text and tabulated information).

Assessment of cost-effectiveness

A systematic review of published economic evaluations of using ELISA tests relative to the alternatives and standard care was undertaken:

- To gain insights into the key drivers of cost-effectiveness of TNF testing.
- To get an overview of the alternative modelling approaches that have been adopted to evaluate the use of therapeutic drug monitoring in people with RA.
- To provide a summary of the findings of previous relevant cost-utility, cost-effectiveness, and cost-benefit studies.

Eligible studies for inclusion to the systematic review were selected according to the following criteria:

- The inclusion criteria for population and interventions were as in the clinical-effectiveness systematic review.
- The following types of economic evaluations were included: cost-utility analyses, cost-effectiveness analyses, cost-benefit analyses, cost-consequence analyses and cost-minimisation analyses. Systematic reviews of economic studies were also included.

Screening was done independently by two reviewers. Disagreements between reviewers were resolved by discussion. All references considered to meet the inclusion criteria by either reviewer at the title and abstract stage were included for full-text screening.

Due to the lack of RCTs identified in the clinical effectiveness systematic review, additional literature searches to identify studies evaluating *any tests* used to monitor anti-TNF-α treatment of people with rheumatoid arthritis were conducted.

The choice of modelling approach was primarily driven by the availability and quality of the evidence identified in the clinical effectiveness systematic review and the additional searches. Due to the limited evidence identified in the clinical-effectiveness systematic review, the multifactorial nature of decisions to adjust treatments in people with RA and the recent changes in the biologics market, which contributed to the uncertainty in the prices of biologics and their uptake within the UK, a simplified modelling approach (threshold analysis), was chosen. Cost-utility analyses were also conducted assuming the estimates of the cost of testing derived from Jani and colleagues (2016). The costs and resource use were considered from the perspective of the NHS and Personal Social Services. No discounting was applied to estimated costs and quality-adjusted life-years (QALYs) due to a short-term time horizon adopted in the study. Clinical outcomes from the INGEBIO study were used in all the analyses. Costs considered in the economic evaluation included the costs of testing, the cost of treatments received by people with RA, and healthcare costs. The costs of testing comprised those of the test kits, staff time to perform the tests and staff training, the cost of the testing service and sample transport. The costs were obtained from the British National Formulary (BNF) and the NHS Reference Costs, from the documents provided by the test manufacturers, and published and unpublished sources. Quality-adjusted life-years (QALYs) were used as as the outcome measure of health benefit. A review of HRQoL studies informed the selection of health states' utilities, and disutilities for flares and AEs. Various sensitivity analyses were conducted to explore the effect of structural and parametric uncertainties on the economic outcomes. Probabilistic sensitivity analysis was deemed inappropriate because of a very substantial variation in clinical practice with respect to treatment, drug dose-tapering and flare management strategies in people with RA. The effect of such variations on the economic outcomes was explored in one-way deterministic sensitivity analyses and a number of scenario analyses.

Results

Clinical effectiveness

No studies met the systematic review's original inclusion criteria.

Eight studies (reported in 11 publications) were included in the systematic review. Three abstracts (Arango and colleagues 2017; Gorostiza and colleagues 2016; Ucar and colleagues 2017) reported the same non-randomised controlled trial (the INGEBIO study),

but this was conducted in people being treated for a range of diseases including RA. The other seven studies were observational, one of which had a historical control group prior to the introduction of treatment monitoring with ELISA tests, and the remaining six studies were single arm studies.

Five of the included studies used Promonitor ELISA kits to monitor drug levels and/or anti-drug antibody levels (INGEBIO; Chen 2016; Inciarte-Mundo 2016, Rosas 2013, and Senabre Gallego 2017), and three studies used Sanquin ELISA kits (Pascual-Salcedo 2013; Paredes 2015; Paredes 2016; Lopez-Casla 2013). No studies were identified evaluating the following ELISA testing kits: IDKmonitor, LISA-TRACKER, RIDASCREEN and MabTrack.

Table 1 summarises which treatments and ELISA kits were used in the studies selected in the clinical-effectiveness systematic review.

Table 1: Clinical effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope

		Promonitor	IDKmonitor	LISA-TRACKER	RIDASCREEN	Sanquin*
ADL	*drug*	✓[1]	X	X	X	✓[2]
	antibody	✓[3]	X[4]	X	X	✓[2]
ETN	*drug*	✓[5]	X	X		✓[2]
	antibody	✓[6]	X	X		✓[2]
IFX	*drug*	✓[7]	X[4]	X	X	✓[8]
	antibody	X	X	X	X	✓[8]
GLM	*drug*	X	X	X		X
	antibody	X	X	X		
CTZ	*drug*			X		X
	antibody			X		

Key:

X Indicates availability of a test to measure drug or antibody level in people treated with the specified TNFi and that no studies have been identified in the clinical-effectiveness systematic review, reporting on using therapeutic drug monitoring for the specified test kit and TNFi.

✓ Indicates availability of a test to measure drug or antibody level in people treated with the specified TNFi and that at least one source for the specified combination of the test kit and TNFi has been identified in the clinical-effectiveness systematic review.

ADL: adalimumab; CTZ: certolizumab pegol; ETN: etanercept; GLM: golimumab; IFX: infliximab; TNFi: tumour necrosis factor inhibitor

Notes:

* The type of Sanquin test kits used in these studies (MabTrack or those used by Sanquin Diagnostic Services) was not reported.

[1] Arango and colleagues 2017, Ucar and colleagues 2017 and Gorostiza and colleagues 2016; Chen and colleagues 2016; Inciarte-Mundo and colleagues 2016, Rosas and colleagues 2013 and Senabre Gallego and colleagues 2017

[2] Paredes and colleagues 2015, Paredes and colleagues 2016 and Pascual-Salcedo and colleagues 2013

[3] Arango and colleagues 2017, Ucar and colleagues 2017 and Gorostiza and colleagues 2016; Chen and colleagues 2016, Rosas and colleagues 2013

[4] Indicates that a test for total anti-drug antibodies is also available (total anti-drug antibodies include both unbound, i.e. free, antibodies and those bound to TNF-α inhibitor)

[5] Inciarte-Mundo and colleagues 2016, Rosas and colleagues 2013, and Senabre Gallego and colleagues 2017

[6] Rosas and colleagues 2013

[7] Inciarte-Mundo and colleagues 2016

[8] Lopez-Casla and colleagues 2013, Paredes and colleagues 2015 and Paredes and colleagues 2016

In studies which used the Sanquin testing service, the type of kits was not reported. No includable evidence on using IDKmonitor, LISA-TRACKER or RIDASCREEN was found. In those studies where antibody testing was conducted, the type of testing (*reflex* or *concurrent* testing) was not specified.

All the selected studies except one reported by Lopez-Casla and colleagues (2013) included individuals in remission or low disease activity (at baseline); Lopez-Casla and colleagues considered a mixed population of people with RA comprising primary and secondary non-responders who had developed clinical inefficacy to IFX.

Two (out of eight) studies included mixed populations, with 37% and 49% of participants with RA in the Arango and colleagues (2017) and Pascual-Salcedo and colleagues (2013), respectively. Moreover, populations considered in the selected studies were relatively small (<70 participants), with the exception being the INGEBIO study which had a (*mixed disease*) population of 169 patients.

The included studies measured drug levels and/or anti-drug antibody levels in people treated with ADL, ETN and/or IFX. There were no studies identified in people treated with CTZ and/or GLM. All but one of the included studies enrolled people with RA who had achieved remission or LDA. Only one study (Lopez-Casla and colleagues 2013) recruited people with RA who had experienced a primary non-response or a secondary non-response. All included studies were judged to be at moderate risk of bias.

Comparative controlled evidence – one study

Three abstracts (Arango and colleagues, 2017; Gorostiza and colleagues, 2016; Ucar and colleagues, 2017) reported the same non-randomised controlled trial (the INGEBIO study), which focused on the population who had achieved treatment target (remission or LDA) and remained clinically stable for at least six months. ADL and anti-ADL antibody levels were measured using Promonitor ELISA kits. Monitoring test results were revealed to physicians in the intervention arm. Physicians did not follow any test-based treatment algorithm for the management of people with RA and they used their best judgements to optimise treatment doses. Such monitoring test results were not revealed to physicians in the control arm. This reflected standard care in Spain where treatment decisions were based on clinical judgements without knowledge of drug levels and anti-drug antibodies.

This trial recruited a mixed disease population of 169 participants, including 63 people with RA. The other participants in the study had psoriatic arthritis (PsA) and ankylosing

spondylitis (AS). The results of the total mixed population were extracted as the study's authors were not willing to provide the separate results for the cohort of participants with RA. Therefore, there is limited generalisability of findings from this mixed disease population to the target RA population of this technology assessment.

The findings from this trial (Ucar and colleagues 2017) showed that, at 18-month follow-up, the rate of flares per patient-year was 0.463 for the intervention group and 0.639 for the control group, with a statistically non-significant rate difference of -0.176 (95% CI -0.379 to 0.0289). There was a non-significant reduction in risk of flare in the intervention group compared with the control group (incidence rate ratio (IRR) 0.7252, 95% CI 0.4997 to 1.0578). Median time to first flare was 145 days for participants in the intervention group and 136.5 days for participants in the control group. The intention-to-treat (ITT) analysis from the abstract by Gorostiza and colleagues (2016) showed that at 34-week follow-up, 67.5% (54/80) in the intervention group remained in remission while 64.0% (32/50) in the control group remained in remission, with the difference in proportions of 3.5% (95% CI -13.3% to 20.3%; p=0.68).

In terms of dose adjustment due to monitoring response, the finding by Arango and colleagues (2017) showed that aADL dose was tapered (i.e. reduced) in 35 participants of the intervention group (35.7%) and in 18 participants of the control group (34.6%). This trial (Ucar and colleagues, 2017) also reported that participants' HRQoL (EQ-5D-5L) measures were higher in the intervention group at all visits compared with the control group (further details were not reported). However, statistically significant results were only observed at Visit 2 (p=0.001) and Visit 3 (p=0.035), and no further details were reported.

Overall, the findings of this controlled study showed that there was a non-significant reduction in risk of flare in the intervention group compared with the control group. HRQoL measures were higher in the intervention group at all visits compared with the control group, with statistically significant results being observed at two visits. However, there was an imbalance at baseline in disease severity between the intervention and control groups, and a lack of adjusting for this baseline imbalance in the analysis of clinical outcomes. Also, there were higher attrition rates for some outcomes. On top of the limited applicability to populations with RA, and the lack of full description of methods (in abstracts), these deficiencies resulted in serious risk of bias associated with the findings.

Evidence from observational studies

Seven observational studies were identified evaluating the effect of TDM (with ELISA-based testing) on clinical outcomes in people with RA who had achieved remission or low disease activity (six studies), or in those who had experienced a primary non-response or a

secondary non-response (one study). The sample size of these observational studies ranged from 36 to 64.

Only one observational study (Pascual-Salcedo and colleagues 2013), of monitoring using Capture ELISA (Sanquin) in people with arthritis receiving IFX, ADL, ETN, had a historical control while the other studies were single-arm cohort studies with no comparator. The observational studies reported changes in disease response (six studies), changes in disease activity (two studies), changes in direction and magnitude of therapeutic dose (three studies), and discontinuation of ineffective therapy (one study). Pascual-Salcedo and colleagues (2013) examined two different time periods, pre- and during-TDM practice. The study showed a non-significant reduction in the mean DAS28 score following the implementation of TDM at seven-year follow-up (pre-TDM: mean 2.51 [SD 0.85] vs.during-TDM: 2.31, [SD 0.52]; p=0.061). This study also showed statistically significant reductions in weekly mean dose per patient by each anti-TNF and increases in mean interval of administration for each anti-TNF following the implementation of TDM; unfortunately, these results related to the wider study sample with more than half of participants with PsA rather than RA.

The findings of the other six, single group, observational studies are not presented in the scientific summary because the weakness of their design does not allow a valid assessment of clinical effectiveness and they did not inform the model-based cost-effectiveness analysis.

Cost effectiveness

Cost-effectiveness systematic review

In the cost-effectiveness systematic review, five studies relevant to the decision problem were found: two were reported as abstracts (Pascual-Salcedo and colleagues, 2013 and Ucar and colleagues, 2017), two as full-text journal article (Krieckaert and colleagues, 2015 and Laine and colleagues, 2016) and a PhD thesis (Gavan, personal communication, 6 August, 2018). Furthermore, only two (out of six) TNF testing kits from the NICE scope (Promonitor and Sanquin) and three (out of five) TNF inhibitors (ADL, ETN, IFX) were considered in the selected studies (Table 2).

Table 2: Cost-effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope

		Promonitor	IDKmonitor	LISA-TRACKER	RIDASCREEN	Sanquin*
ADL	*drug*	✓[1]	X	X	X	✓[2]
	antibody	✓[3]	X[4]	X	X	✓[5]
ETN	*drug*	✓[6]	X	X		X
	antibody	X	X	X		X
IFX	*drug*	✓[6]	X[4]	X	X	X
	antibody	X	X	X	X	✓[6]
GLM	*drug*	X	X	X		X
	antibody	X	X	X		
CTZ	*drug*			X		X
	antibody			X		

Notes:

X Indicates availability of a test to measure drug or antibody level in people treated with the specified TNFi and that no studies have been identified in the clinical-effectiveness systematic review, reporting on using therapeutic drug monitoring for the specified test kit and TNFi.

✓ Indicates availability of a test to measure drug or antibody level in people treated with the specified TNFi and that at least one source for the specified combination of the test kit and TNFi has been identified in the cost-effectiveness systematic reviews.

ADL: adalimumab; CTZ: certolizumab pegol; ETN: etanercept; GLM: golimumab; IFX: infliximab; TNFi: tumour necrosis factor inhibitor

Notes:

* The type of Sanquin test kits used in these studies (MabTrack or those used by Sanquin Diagnostic Services) was not reported.

[1] Gavan (2017), Laine and colleagues (2016) and Ucar and colleagues 2017

[2] Krieckaert and colleagues (2015) [1] and Laine and colleagues (2016) [2]

[3] Laine and colleagues (2016) [2] and Ucar and colleagues 2017

[4] Indicates that a test for total anti-drug antibodies is also available (total anti-drug antibodies include both unbound, i.e. free, antibodies and those bound to TNF-α inhibitor).

[5] Gavan (2017) (personal communication) and Laine and colleagues (2016)

[6] Laine and colleagues (2016)

Three (out of five) studies (Krieckaert and colleagues, 2015, Laine and colleagues, 2016 and Gavan, 2017) were model-based; the modelling approaches employed in those studies were Markov (Krieckaert and colleagues, 2015 and Laine and colleagues, 2016) and discrete-event simulation (DES) (Gavan, 2017). The other two studies (Ucar and colleagues 2017 and Pascual-Salcedo and colleagues 2013) were cohort-based.

Krieckaert and colleagues (2015) and Laine and colleagues (2016) used patient-level data for the parameterisation of their Markov models to compare TDM with standard care. The analysis in Krieckaert and colleagues (2015) incorporated data on 272 people with RA ptreated with ADL, and data on direct medical costs and health-related quality of life (HRQoL) from the Utrecht Rheumatoid Arthritis Cohort (URAC) study group (N=1,034). Laine and colleagues (2016) based their analysis on patient-level data from 486 and 1,137

people treated with ADA and IFX, respectively, from the clinical sample registry of United Medix Laboratories Ltd (Finland).

The DES model developed by Gavan was parameterised from various clinical effectiveness sources and published literature on cost-effectiveness of biologic therapies in people with RA, including summary data from the British Society for Rheumatology Biologics Register (BSRBR) such as the mean and standard deviation for patient age and HAQ score, and gender composition.

Findings in included studies

Both Krieckaert and colleagues (2015) and Laine and colleagues (2016) reported that TDM was cost saving compared to standard care, based on follow-up periods of up to three years. The Krieckaert and colleagues (2015) study, in the Netherlands, reported a formal cost per QALY analysis in which TDM dominated standard care in the base-case scenario in 72% of simulations. The incremental cost-effectiveness ratios (ICERs) are arguably somewhat meaningless given the small QALY differentials involved. In a range of sensitivity analyses, a net loss of QALYs with the intervention was associated with drug level cut-offs, the use of EULAR good response as an outcome, or the use of non-TNF inhibiting biologicals. With regard to UK clinical practice, Krieckaert and colleagues (2015) modelled testing at the 28th week, and considered dose reduction by prolongation of the interval between drug administrations in responders with high levels of adalimumab (DL>12mg/l), which may differ from UK clinical practice.

The Laine and colleagues (2016) study, conducted in Finland, did not report a cost per QALY analysis, although the authors attempted to analyse the frequency and cost impact of not using TDM with regard to inappropriate treatment decisions (e.g. continuation of ineffective therapy). The assumption was made that participants in the routine practice arm would typically experience three months' delay in receiving optimal treatment compared to participants in the intervention arm. This was justified based on the typical follow-up intervals of participants in Finland. Also, in Finland anti-drug antibody levels of at least 30 U/mL rather than 12 U/mL are considered clinically significant.

Both the INGEBIO study (Arango and colleagues, 2017, n=169 patients) and Pascual-Salcedo and colleagues (2013) (n=88 patients) recruited mixed populations featuring respectively 37% and 49% of subjects with a diagnosis of RA. In addition, there were only limited details of the input parameters and analysis, specifically:

- No details of utility values or incremental QALY outcomes were provided.

- The studies did not consider specific test-based treatment algorithms.
- Pascual-Salcedo and colleagues did not specify which ELISA test kits were used in their study.

Although the INGEBIO study was the only controlled effectiveness study, it had other limitations which have already been described in the clinical effectiveness section

The recent study by Gavan (2017) perhaps most closely matches the decision problem. In this study, modelling was based on data from the BSRBR register which is the main source of evidence on the use of biologics in people with RA in the UK. Furthermore, the research questions addressed in Gavan (2017) are most relevant to the decision problem considered in this report. Gavan (2017), however, did not consider any specific test kit, and only ADL treatment was modelled as first line.

Of Gavan's three stated research questions (Gavan, 2017 p. 59), namely:

- Research Question 1: *What was the existing economic evidence for stratified medicine in RA?*
- Research Question 2: *How were treatment decisions with biologic therapies made for patients with RA in current practice in England?*
- Research Question 3: *Are treatment decisions stratified by ADL antibodies and drug level testing, for patients with RA in England, a relatively cost-effective use of health care resources?*

Question 3 aligns closely with the decision problem. The analysis of the strategy of dose reduction after two years for people in remission suggests that testing may not be cost-effective. The net monetary benefit per patient relative to standard care was of £3,196.72, with an associated mean QALY loss of -0.000121 per patient. However, Gavan (2017) pointed out that there was a large element of decision uncertainty. The decision uncertainty was mostly driven by uncertainty in the cost of testing and test accuracy.

Considerations in the development of the independent economic assessment

There several possible factors that may drive either the health impacts/benefits and/or additional costs or cost savings associated with therapeutic drug monitoring in people having therapeutic drug monitoring when treated with bDMARDs. The main ones are:

- Cost of testing
- Concurrent vs reflex testing
- Singlet versus duplicate testing

- Costs of drugs: originators vs biosimilars
- Tapering (dose reduction)
- Outcomes: risk of flares
- Outcomes: rate of remission

Cost of testing

It is self-evident that the cost of testing will be a driver of TDM. Estimates of the cost of testing in the UK are available from both a published study (using Promonitor kits, Jani and colleagues (2016)) and from the Exeter Clinical Laboratory (Royal Devon and Exeter NHS Foundation Trust) which conducts approximately 80% of testing for monitoring biologics in the UK. In the economic analysis, the estimates from Jani and colleagues (2016) were used. The cost at the RD&E provided by Dr Timothy McDonald was based on the costs of IDKmonitor test kits, and therefore was not utilized in our economic analysis of Promonitor and ADL. The costs of assays were derived from information submitted by the manufacturers of the test kits.

Concurrent versus reflex testing

Concurrent testing of both drug levels and antibody levels in the blood might be more expensive than reflex testing – in which antibody testing only occurs when drug levels are undetectable – but this is not certain, because reflex testing may require a new blood sample to be obtained plus the (negligible) cost of phone calls to request antibody testing.

Singlet versus duplicate testing

Performing ELISAs once per person (*singlet* testing) incurs a lower cost compared to *duplicate* testing (ELISA twice per patient). However, it is less precise. Therefore, *duplicate* testing was selected in the base case analysis conducted by Jani and colleagues in the microcosting study (Jani and colleagues, 2016). However, based on clinical advice, *singlet* testing is more common in the UK (Dr Timothy McDonald, personal communication, December 2018). Therefore, we adopted this approach in our primary analysis, and we conducted an additional analysis assuming *duplicate* testing.

Frequency of TNF testing

More frequent TNF treatment monitoring will be more costly in terms of the cost of testing. There is no recommended protocol for how frequently TNF treatment monitoring should be conducted, but two sources suggest it might be approximately once a year. Rosas and colleagues (2015) reported the total number of drug and antibody monitoring tests in people with RA in remission over a two-year period (94 tests in 45 patients), which is about *one test*

per patient per year. Dr Meghna Jani (clinical advisor) also confirmed that in people in remission/under routine follow up, TNF testing may be conducted up to once a year.

Costs of originator drugs and their biosimilars

The biological medicines market will increase in complexity over the coming months and years as more originator biological medicines lose patent exclusivity and additional biosimilar medicines come to market. The patent for one of the TNF inhibitors, ADL (known by brand name Humira®), expired on 16 October, 2018. New medications with similar active properties ("biosimilar" versions) will become available in the NHS in the end of 2018. The following ADL biosimilars have already been approved for use in the UK but not yet launched (as of 30 November, 2018):

- Amgevita® (Amgen)
- Hulio® (Mylan/Fujifilm Kyowa Kirin)
- Hyrimoz® (Sandoz)
- Imraldi® (Samsung Biogen)

According to Regional Medicines Optimisation Committee Briefing, at least two further biosimilars are expected to become available in the UK during 2019: Cyltezo® (from Boehringer Ingelheim) and the second will be brought to the market by Fresenius Kabi.

As for the current uptake of biosimilars in the UK, according to the Medicines Optimisation Dashboard Data published by the NHS England (September 2018 release), 92% and 85% of people who were prescribed IFX and ETN, respectively, are taking biosimilars. However, there are regional variations in the uptake of different biosimilars since the cost per dose is negotiated at the Clinical Commissioning Group (CCG) level.

Although the NICE guidance recommends that people with RA receive the anti-TNFs with the lowest acquisition and administration costs, in practice, however, other non-cost factors such as individual characteristics, hospital characteristics and changes in regional rheumatology clinical guidelines may influence treatment selection (Gavan 2017).

Flares

The concept of flare remains challenging to understand as there are no generally recognised definitions of or well-validated measures for flare in RA (Bykerk and colleagues, 2014). Nevertheless this term referring to episodes of worsening disease activity which includes a range of symptoms of different duration and magnitude is commonly used (Bingham and colleagues, 2009[333]).

There is a substantial heterogeneity in the duration of flare. Based on clinical advice, it may vary from two to three days to two to three months depending on severity, with more severe flares requiring specialist treatment

Tapering

According to EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs, tapering of bDMARDs should be considered in people in persistent remission after having tapered glucocorticosteroids (GC), especially if this treatment is combined with a conventional synthetic DMARD. Tapering here means reduction of dose (e.g. reducing etanercept 50mg to 25mg/ week (Smolen and colleagues, 2013) or extension of interval between applications, *spacing* (e.g. increasing the interval between ADL injections to 10 days rather than one week as in the Exeter Biologic Clinic recommendations described in Appendix 5).

Currently there is no gold standard in the UK on how dose tapering should be performed. Studies evaluating dose tapering have used different approaches. In clinical practice, dose tapering varies extensively depending on clinical opinion.

Modelling approach

Due to the limited evidence identified in the clinical-effectiveness systematic review, the multifactorial nature of decisions to adjust treatments in RA patients and the recent changes in the market of biologics, which contributed to the uncertainty in the prices of biologics and their uptake within the UK, a simplified modelling approach – a threshold analysis – was the main form of cost-effectiveness analysis used. Although not a formal cost-effectiveness analysis, this approach allows the estimation of the cost of TNF testing at which the net monetary benefit (NMB) of the ELISA test-based monitoring becomes zero, while taking into consideration the major components of differential costs and QALYs.

Clinical outcomes from the INGEBIO study (Ucar and colleagues 2017 and Arango and colleagues 2017) were used to inform assumptions and parameters in all analyses. The differential costs of drug acquisition, drug administration, and disease management were included in the model; the latter comprised the costs of managing flares and adverse events, and the costs associated with managing different health states. QALYs were estimated from the rates of flares and adverse events, and the duration of remission and low disease activity (LDA)/active disease health states in the intervention and control arms.

In addition to the threshold analysis, ICERs were estimated using the list prices of the originator products and their biosimilars assuming similar clinical effectiveness across the

anti-TNFs and similar performance of the different test kits. Estimates of the cost of testing were based on Jani and colleagues (2016) and clinical advice.

Exploratory analysis for those in remission/LDA

The clinical and economic effect of ADL tapering in people with RA in remission/LDA was evaluated in the INGEBIO study (Arango and colleagues, 2017, Ucar and colleagues, 2017 and Gorostiza and colleagues, 2016).

The clinical outcomes reported in Ucar and colleagues (2017) were incorporated in our economic analysis to estimate the cost of drug and antibody testing, at which the addition of ELISA testing to usual clinical practice would result in zero NMB.

Since the patent for the adalimumab originator product, Humira®, has expired in October 2018, and the acquisition costs for the ADL biosimilars were not known to the AG at the time of writing, in the threshold analysis, the annual acquisition cost was varied from £1,000 to £9,187 per patient-year; the latter represents the annual cost of Humira® assuming the dose of 40 mg every two weeks delivered by subcutaneous injection using pre-filled pen and the NHS indicative price from the BNF.

The other major assumptions were as follows, with further details in Table 4:

- ADL dose tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).
- Dose is tapered in a proportion of people in each arm at the start of simulation.
- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored *in all people* on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), and, therefore, the administration cost is zero.
- The costs associated with flare management are:
    - £423 per flare for diagnostic investigations
    - £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.

- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.
- The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
- The duration of AE is 28 days.
- The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 3: Model assumptions

Assumption	Estimate	Source	Relevant table/ sections in the report
Dose tapering strategy	Spacing: from 40 mg of ADL every 2 weeks to 40 mg every 3 weeks	1st dose reduction in the Exeter biologic clinic recommendations (Appendix 5)	Section 4.1.9.1.5
Proportion of patients on tapered dose:			
Intervention	35.8%		Table 40
Control	36.7%		Table 40
Proportion of flared patients in whom the full dose is restored	100%	Exeter biologic clinic recommendations	Appendix 5
Mean duration of remission (days)			
Intervention	344		Table 40
Control	329		Table 40
Mean follow-up (days)	505	As in the control arm (Ucar and colleagues, 2017)	Table 40
Acquisition costs (per patient-year): Humira®			
Full dose[1]	£9,187	BNF	Section 4.1.9.1.3
Tapered dose	£6,125	BNF, Exeter biologic clinic recommendations	Appendix 5
Flared patients[2]	£9,187	BNF, Exeter biologic clinic recommendations	Appendix 5
Treatment wastage on full dose (per patient-year)	£370	Clinical advice	Section 4.1.9.1.6

Assumption	Estimate	Source	Relevant table/ sections in the report
Administration cost for Humira® (ADL) (per patient-year)[2]	£0	Clinical advice	Section 4.1.9.1.7
Cost of flare management[3, 4]	£423/per flare	Cost of diagnostic investigations (Maravic and colleagues, 2005[4])	Section 4.1.9.1.19
	£68/month	Monthly cost of treatment (excluding DMARDs) (Maravic and colleagues, 2005[4])	Section 4.1.9.1.19
Cost of managing health states (per patient-year)[5]			
Remission	£11,409	Barbieri and colleagues (2005),[5] Radner and colleagues (2014),[6]	Section 4.1.9.1.16
LDA/active disease	£18,889	National Schedule of Reference Costs 2017-18[7]	Section 4.1.9.1.16
Cost of managing AEs (per infection)	£1,622[6]	TA375[8]	Section 4.1.9.1.20
Utilities			
Remission	0.717	Estimated from HAQ scores for different HAQ bands reported by Radner and colleagues (2014)[6]	Section 4.1.9.2.1
LDA/active disease	0.586[7]		Section 4.1.9.2.1
Disutility of flare	0.140	Markusse and colleagues, 2015[9]	Section 4.1.9.2.2
Disutility of AEs	0.156	TA375,[8] Oppong and colleagues (2013)[10]	Section 4.1.9.2.3
Flare rate			
Intervention	0.463	Ucar and colleagues 2017[11]	Section 4.1.8.1.1
Control	0.639	Ucar and colleagues 2017[11]	Section 4.1.8.1.1
Mean time to first flare (days)			
Intervention	208.07	Derived from Kaplan-Meier estimates (from the INGEBIO study) of time to first flare, provided by Ucar and colleagues (personal communication, 9 September, 2018)	Section 4.1.8.1.3
Control	189.32		Section 4.1.8.1.3
Flare duration (days)[8]	7	TA375[8]	Section 4.1.8.1.2

Assumption	Estimate	Source	Relevant table/ sections in the report
Rate of AEs			
Patients on full ADL dose	3/100 patient-years	Senabre Gallego and colleagues (2017)[12]	Section 4.1.8.2.1
Patients on reduced ADL dose	2/100 patient-years[9]	Singh and colleagues (2015)[13]	Section 4.1.8.2.1
Duration of AE (days)	28	TA375,[8] Oppong and colleagues (2013)[10]	Section 4.1.8.2.2

Key: ADL: adalimumab; AE: adverse event; BNF: British national Formulary; HAQ: health assessment questionnaire; HAD: high disease activity; MDA: moderate disease activity; OR: odds ratio; PPP: purchasing power parities; RA: rheumatoid arthritis; RCTs: randomized controlled trials; TA: technology appraisal

Notes:

[1] Assuming 40 mg every two weeks by subcutaneous injection using pre-filled pen, and NHS indicative price from the BNF.

[2] The mean time to first flare was estimated from additional evidence (Kaplan-Meier curves for time to first flare) from the INGEBIO study provided by Ucar and colleagues (2007)[11] (poster, personal communication).

[3] The estimates were derived from the costs of managing a flare in a hypothetical person with a 10-year history of RA in the French setting. The costs were converted to pound sterling based on PPP and inflated to 2017-18 prices using the healthcare price index (Section 4.1.9.1.1).

[4] The estimates from Maravic and colleagues (2005)[4] do not include the cost of rheumatology appointments.

[5] The costs of managing health states were included by HAQ-dependency, i.e. by assigning an annual cost to mutually exclusive HAQ intervals.

[6] The estimate of £1,479 per patient-year from the source was inflated to 2017-18 prices using the healthcare price index (Section 4.1.9.1.11).

[7] The estimate was computed from a weighted average HAQ score for the LDA, MDA, HDA health states reported by Radner and colleagues (2014)[6] and mapped to EQ-5D values following Malottki and colleagues (2011)[14] (Section 4.1.9.2.1).

[8] This estimate was used for calculation of QALYs only since it was assumed that the ADL dose in people with flares is switched back to the full dose indefinitely.

[9] Based on OR=1.31 for standard-dose biologics in people with RA reported by Singh and colleagues (2015).[13] The OR estimate was obtained in a Bayesian network meta-analysis (using a binomial likelihood model) of 11 published RCTs (n=4,788) to assess the risk of serious infections in anti-TNF-biologic-experienced people with RA.

In the primary analysis, QALYs were estimated based on the heath-state utilities for remission and LDA/active disease, and with short-term disutilities associated with flares and AEs (i.e. as adjustments to these underlying chronic health states). It was assumed that individuals in any health state (i.e. in remission, LDA and active disease) can experience flares. Utilities for the mixed disease population in the INGEBIO study were assumed to be the same as those for the population of people with RA, since no evidence on HRQoL directly relevant to the population considered in INGEBIO has been identified. These utilities were calculated using UK tariffs.

When modelling the effect of AEs on HRQoL and costs, the AG adopted the approach used in a previous technology assessment for NICE (TA375) - it was assumed that only serious adverse events (serious infections in particular) would carry a significant cost and disutility burden.

Mortality associated with RA was not modelled, and no discounting was applied to the costs and outcomes due to the short-term time horizon of about 18 months adopted in this study.

All costs were inflated to 2017-18 prices using the healthcare price index.

Results: adalimumab and Promonitor

Threshold analysis

The results of the threshold analysis, assuming the Promonitor test kit is used to monitor people with RA in remission/LDA recieving originator ADL (Humira®) are presented in Table 4 and Figure 1. Figure 1 shows the annual cost of ELISA-based testing at which TDM would become cost-effective at the two WTP thresholds used in NICE decision making for the range of ADL acquisition costs of £1,000–£9,187. Since the data reported in Arango and colleagues (2017)[15] are for a longer follow-up than that reported in Ucar and colleagues (2017), the results using the two different reports of the outcomes of the INGEBIO study are presented.

If the results of Ucar and colleagues (2017) are used, then with the current price of originator ADL, testing would need to be cheaper than £391 per year in order for TDM to be judged as cost-effective. Using the results presented in Arango and colleagues (2017), however, there would be no cost of testing at which testing becomes cost-effective (because using these outcomes testing was estimated to be both more costly and less effective than standard care).

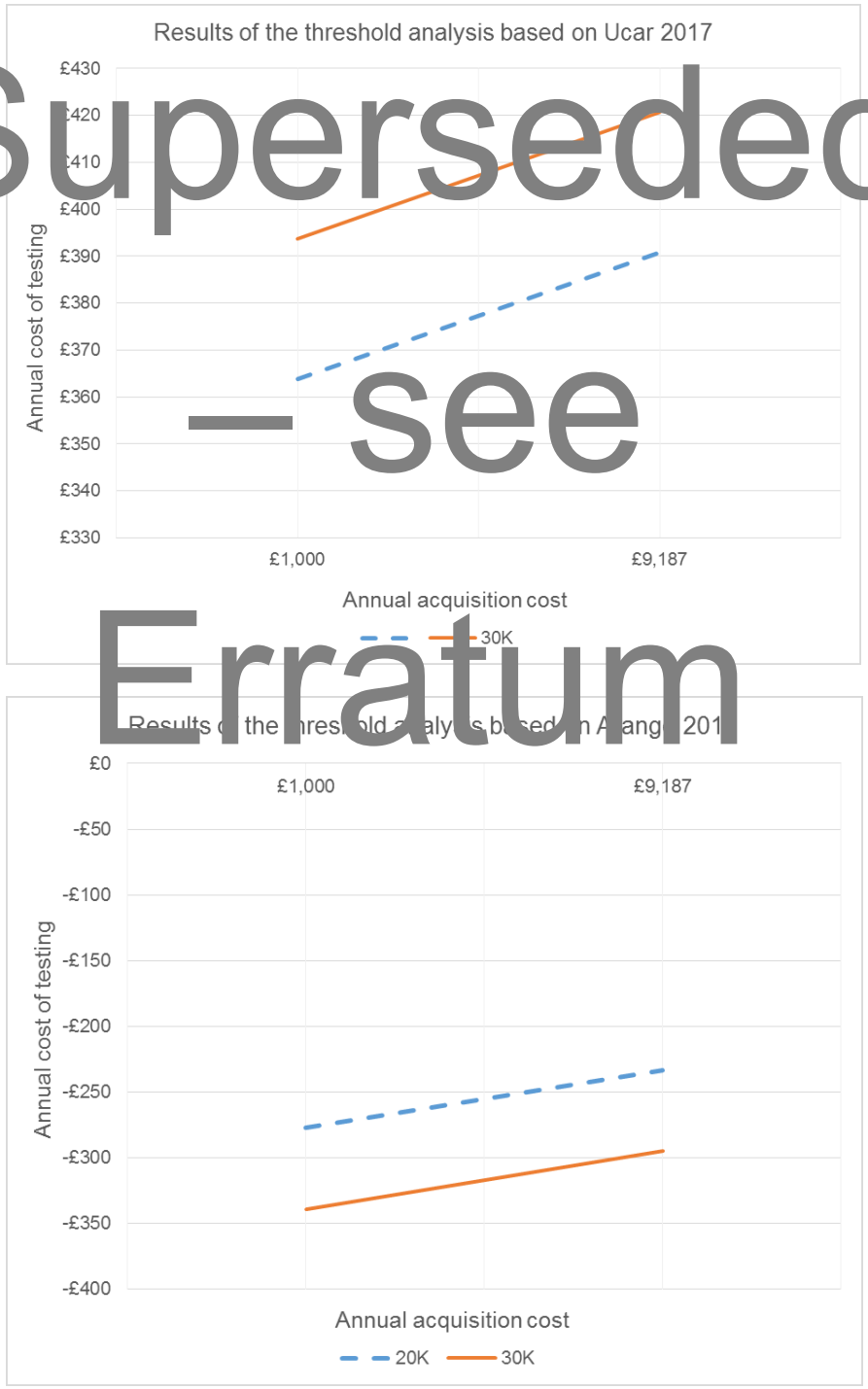
Table 4: Threshold value for the cost of testing

ICER threshold	Results based on INGEBIO study, Ucar and colleagues 2017	Results based on INGEBIO study, Arango and colleagues 2017
£20,000	£391	-£233
£30,000	£421	-£295

Key: ICER: incremental cost-effectiveness ratio

Such differences in the results are due to differences in the mean duration of remission (as reported in Ucar 2017) and remission/LDA (Arango 2017) between the control and intervention arms. Arango reported a longer duration of remission/LDA in the control group than in the intervention group (475.2 versus 460.2 days), while Ucar and colleagues 2017 reported a longer duration in the intervention group (344 versus 329 days in the control group).

Figure 1: Results of the threshold analyses Arango and colleagues (2017) and Ucar and colleagues (2017)



Source: Ucar and colleagues (2017) and Arango and colleagues (2017)

These results are inconclusive for two reasons. First, because they are in opposite directions and, second because they are based on very small and uncertain differences in outcomes (QALY differences of less than 0.01). The negative value of the cost of testing at which NMB equals zero means that, when using the trial results as presented in Arango and

colleagues (2017), there are no (positive) values of the cost of testing at which it would be a cost-effective option.

Cost-utility analysis

The incremental QALYs and incremental costs for testing versus standard care strategy are shown in Table 5, assuming:

- regular testing is undertaken in people with RA in remission/LDA treated with Humira® and tested using Promonitor
- the costs of testing are as in Jani and colleagues (2016)
- the frequency of testing is *one test per patient-year* and
- that testing of drug and antibody levels is done *concurrently (singlet dilution) in a UK laboratory.*

The outcome data were derived from two reports of the IMAGINE BIO study, Ucar 2017 and Arango 2017.

Table 5: Cost-effectiveness results in patients in remission/LDA treated with Humira® and tested using Promonitor

	Intervention arm	Control arm	Intervention vs. control
Based on Ucar and colleagues (2017)			
QALYs (mean)	1.108	1.103	0.004
Total costs (mean)	£32,178	£32,438	-£260
ICER (Cost / QALY gained)			ICER not relevant - Intervention dominates standard care
Based on Arango and colleagues (2017)			
QALYs (mean,)	1.138	1.147	-0.009
Total costs (mean)	£36,284	£35,923	£361
ICER (Cost / QALY gained)			ICER not relevant - Standard care dominates Intervention

Key: ICER: incremental cost effectiveness ratio; QALY: quality-adjusted life year; vs: versus
Note: The postage was £4 per parcel.

As with the threshold analyses, these results are inconclusive for two reasons. First, because they are in opposite directions and, second because they are based on very small and uncertain differences in outcomes (QALY differences of less than 0.01). Furthermore, it is not possible to argue that either the analysis based of Ucar and colleagues (2017) or that based on Arango and colleagues (2017) is more valid than the other – they both have significant weaknesses. The follow-up in Arango and colleagues (2017) is over a longer time

horizon (545 days in the control arm) than Ucar and colleagues (2017) (505 days in the control arm). The fact that these different analyses from the same study produce opposite estimates of effects and costs further highlights the uncertainty, which the economic analysis for this appraisal may only serve to amplify.

A number of additional sensitivity analyses were undertaken to explore the impact of parametric and structural uncertainty on the outcomes reported in Table 5. The results of the sensitivity analyses are shown in Table 6.

Table 6: Additional sensitivity analyses (people in remission/low disease activity)

Sensitivity analysis	Assumptions	ICER		Source (relevant sections)
		Ucar and colleagues (2017)	Arango and colleagues (2017)	
Impact of flares only (health states and AEs are not included)	Only flares contribute to differential costs and QALYs	ICER not relevant - Standard care dominates Intervention	ICER not relevant - Standard care dominates Intervention	Scenario C (people in remission, Gavan 2017, Section 3.3.2.3)
Tapering strategy	Spacing: reduction of ADA dose to 40mg every 4 weeks	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	2nd dose reduction Exeter biologic clinic recommendations (Appendix 5)
Treatment wastage	No wastage	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Assumption
Flare duration, days	19	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Weighted average based on Bykerk and colleagues (2014) [16 16 16 16 16 16] and clinical advice
Proportion of flared patients in whom full dose is restored	55%	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Bykerk and colleagues (2014) and clinical advice
	0%	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Clinical advice
Utilities[2]				
Remission	0.496	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Estimated from HAQ scores reported in TA375 (Fig. 94, p.366) (Section 4.1.9.2.1)
LDA/active disease	0.302			

Sensitivity analysis	Assumptions	ICER		Source (relevant sections)
		Ucar and colleagues (2017)	**Arango and colleagues (2017)**	
Disutility of flare	0.085	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Minor B type of utility (Table 69, Section 4.1.9.2.2)
	0.116	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Major B type of utility, (Table 69, Section 4.1.9.2.2)
Frequency of testing (tests/year)	2	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Rosas and colleagues 2015, clinical advice (Section 4.1.9.1.20)
Cost of testing	Refer to Table 56 for the cost of testing			
Duplicate concurrent testing with initial phlebotomy appointment		In all analyses, ICER not relevant - Intervention dominates standard care	In all analyses, ICER not relevant - Standard care dominates Intervention	Jani and colleagues (2015) (Section 4.1.9.1.14)
Duplicate reflex testing without initial phlebotomy appointment, 35.8% of ptxs w/LDL[3,4]				
Duplicate reflex testing with initial phlebotomy appointment, 35.8% of ptxs w/LDL[4]				
Singlet reflex testing without initial phlebotomy appointment, 35.8% of ptxs w/LDL[3,4]				
Singlet reflex testing with initial appointment, 35.8% of ptxs w/LDL[4]				

Sensitivity analysis	Assumptions	ICER		Source (relevant sections)
		Ucar and colleagues (2017)	Arango and colleagues (2017)	
Duplicate concurrent testing without initial phlebotomy appointment[3]				
Duplicate reflex testing without initial phlebotomy appointment, 4.7% of ptxs w/LDL[3,5]				
Duplicate reflex testing with initial phlebotomy appointment, 4.7% of ptxs w/LDL[3,5]				
Singlet concurrent testing without initial phlebotomy appointment[3]				
Singlet reflex testing without initial phlebotomy appointment, 4.7% of ptxs w/LDL[3,5]				
Singlet reflex testing with initial appointment, 4.7% of ptxs w/LDL[5]				

Key: AE: adverse events; HAQ: health assessment questionnaire; ICER, incremental cost-effectiveness ratio; LDA: low disease activity; LDL: low drug level

Notes:

All costs are reported in 2017-18 prices.

[1] Based on the average cost of joint replacement surgery in rheumatoid arthritis patients from the Royal Devon & Exeter NHS Foundation Trust (Appendix 8).

[2] Utilities for the mixed disease population (as in the INGEBIO study) were assumed to be the same as those for people with RA

[3] The cost of testing does not include the cost of an additional phlebotomy appointment which might not be required if people will receive regular hematological analysis as part of on-going treatement.

[4] Assuming 35.8% of people have low drug level (Laine and colleagues 2016)

[5] Assuming 4.7% of people have low drug level (Chen and colleagues 2015))

In all but one sensitivity analysis based on Ucar and colleagues (2017), the intervention dominated standard care. When the impact of *flares only* was modelled (i.e. health states and AEs were not included), standard care dominated test-based treatment strategy (Table 6). In none of the sensitivity analyses based on data from Arango and colleagues (2017) was there any change to the finding that standard care dominates the intervention.

One-way sensitivity analyses for some of the parameters used to estimate the ICERs based on data from Arango and colleagues (2017) were also conducted (Table 7). Changing these parameters had no impact on the findings; standard care was estimated to dominate the intervention in all analyses.

Table 7: One-way deterministic sensitivity analyses

Parameter	Assumption	ICER	Source
Percentage of people in whom the biologic was tapered	+20% in the intervention arm and -20% in the control arm	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017)
Flare rate	-20% in the intervention arm, +20% in the control arm	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017)
Differential time in remission	+10% in the intervention arm, -10% in the control arm of the differential time in remission	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017)
Costs of managing health states	- 20%	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017), Radner and colleagues (2014), Barbieri and colleagues (2005)

Key: ICER: incremental cost-effectiveness ratio

Sensitivity analyses were conducted and are presented in the main cost-effectiveness results section of the report.

Other scenario analyses considered but not conducted due to no or low quality clinical data were: analysis of testing in the context of primary or secondary non-response; analysis for non-responders who did not adhere to treatment with biologic therapies, including switching to intravenously administered IFX.

Probabilistic sensitivity analysis was deemed inappropriate because of a very substantial variation in clinical practice with respect to disease management in people with RA in England.

Results: etanercept and Infliximab and Promonitor

The cost-effectiveness of TNF testing in people treated with Etanercept (originator and biosimilar) and Infliximab (biosimilar) using the Promonitor test kits was explored in scenario analyses. Enbrel is the most expensive originator product considered in this assessment, while biosimilars Erelzi and Flixabi/Renflexis have the lowest acquisition costs among the TNF inhibitors administered via subcutaneous and intravenous routes, respectively.

In those analyses, it was assumed, based on clinical advice (and a lack of evidence to the contrary), that the clinical effectiveness of the different TNF inhibitors is likely to be the same, and the clinical effectiveness estimates from Ucar and colleagues (2017) were adopted, with all assumptions except acquisition and administration costs, as in Table 3.

The information on the actual costs to the NHS of the TNF inhibitors was not available to the AG at the time of writing, and therefore the list prices of the biologics were assumed. The results are presented in Table 8.

Table 8: Cost-effectiveness results for the other tests and TNF inhibitors: people in remission/LDA

Treatment		ICER	
	Cost per year (£)	**Ucar and colleagues (2017)**	**Arango and colleagues (2017)**
Etanercept			
*Enbrel®**	9,327	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention
Erelzi	8,394	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention
Infliximab[1]			
Flixabi/Renflexis (no wastage)	5164	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention

Key: ETN: etanercept; ICER: incremental cost-effectiveness ratio; IFX: infliximab

Notes:

For all test kits except those used by Sanquin, it was assummed that blood samples would be sent for testing to UK laboratories, and the postage of £4 (per small parcel) was applied; in the scenarios modelling MabTrack and Sanquin Diagnostic Services, the postage of £10 (per small parcel sent to Netherlands) was modelled [17].

* The originator (or reference) product

[1] IFX administration cost was assumed to be 283 per injection (Section 4.1.9.1.7).

Discussion

Strengths and limitations

Clinical effectiveness

A comprehensive literature search was undertaken to identify both published and unpublished studies. These included electronic searches of a variety of bibliographic databases as well as screening of clinical trial registers and conference proceedings to identify unpublished studies. The review process followed recommended methods to minimise the potential for error and/or bias. The quality of included studies was assessed and accounted for when interpreting the review results. Appropriate synthesis methods were employed taking into account the heterogeneity of study characteristics.

In terms of limitations, non-English-language studies were excluded. Few relevant studies were identified for the evaluation of clinical effectiveness of TDM based on ELISA testing in the target populations. Evidence was particularly scarce in relation to clinical effectiveness of therapeutic drug monitoring in people with RA who had experienced a primary non-response or a secondary non-response. Only one relevant non-randomised controlled trial that was conducted in a mixed disease population (including only 37% people with RA) was identified. No studies were identified evaluating the following ELISA kits: IDKmonitor, LISA-TRACKER, RIDASCREEN and MabTrack. There was considerable clinical heterogeneity associated with interventions, outcomes and length of follow-up between included studies.

Only in the INGEBIO study, also included in the systematic review of cost-effectiveness studies, was an ELISA test-guided treatment compared against standard care/monitoring. In this study, however, physicians were not obliged to follow any test-based treatment algorithm but could use testing to alter doses, based on their judgement, in patients from the intervention arm. Moreover, the study was reported in abstracts only, and the reported outcomes may not be directly relevant to the NHS clinical practice since the study was conducted in Spain. Therefore, an additional systematic literature review to identify RCTs evaluating *any tests* used to monitor anti-TNF-α treatment of people with RA was conducted to support the economic assessment. However, no relevant sources were identified.

Cost effectiveness – systematic review evidence

A systematic review of published economic evaluations of using ELISA tests relative to the alternatives and standard care was undertaken to help inform the type and structure of the independent economic assessment. The results of this review indicate limited existing evidence on the cost-effectiveness of therapeutic drug monitoring in people with RA. Despite

a comprehensive search of the literature, only five studies were identified. Two (out of five) TNF testing kits from the NICE scope (Promonitor and Sanquin) and three (out of five) TNF inhibitors (ADL, ETN and IFX) have been assessed in the selected studies. The systematic review was also limited by reporting as two (out of five) selected studies were reported as abstracts. These studies therefore mainly informed the planning of the independent model-based analysis.

Cost effectiveness – model-based analysis

Despite substantial weaknesses in the clinical effectiveness evidence base, a simple model was developed to estimate the cost-utility of ELISA test-based monitoring for people with RA taking bDMARDS.

The analyses conducted are inconclusive and suggest considerable uncertainty in the cost-effetciveness of therapeutic monitoring of TNF-alpha inhibitors in RA. Data from two reports of the same study produced very different conclusions on the cost-effectiveness of Promonitor testing in people receiving ADL who are in remission/LDA. The results based on the longer follow-up (Arango and colleagues 2017[15]) suggested that monitoring is more costly and produces fewer QALYs than standard care.

Of the sensitivity analyses conducted, only the assumption that the rate of flares alone changes as a consequence of monitoring, impacted on the results. This was when evidence from Ucar and colleagues (2017) was used and resulted in standard care dominating the intervention.

Exploratory analyses of using Promonitor to monitor patients in remission/LDA receiving ETN or IFX were undertaken, and showed the same results as that for ADL: using the longer follow-up (Arango and colleagues 2017) monitoring is more costly and produces fewer QALYs than standard care.

The main effectiveness evidence in the model was from the poorly reported INGEBIO study (a non-randomised controlled trial from Spain, where <40% of participants had RA), heavily supplemented by input parameters from other studies and expert advice. The results of the economic analysis should therefore be viewed as exploratory and highly speculative. For example, although the INGEBIO study only evaluated testing using Promonitor ELISA kits, for those in remission/LDA treated with Humira® (ADL), with further assumptions these results have been used to estimate the threshold testing costs at which TDM would become cost-effective with people taking other TNF inhibitors (and taking either originator products or biosimilars.

In summary, there is much uncertainty in relation to key potential drivers of the effectiveness and cost-effectiveness of using ELISA based testing to monitoring treatment with bDMARDs in people with RA, that no firm conclusions can be drawn.

The most important limitations of the independent economic analysis are:

- Limited evidence from comparative studies on clinical effectiveness, health-related quality of life (HRQoL) and costs associated with test-based treatment strategies. Due to the paucity of data, not all test kits and TNF inhibitors, and not all patient populations specified in the NICE scope were considered in the primary economic analysis. There was no sufficiently valid and reliable evidence related to primary non-responders and secondary non-responders. Moreover, no economic evaluations relevant to IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits have been conducted.
- Several test–based treatment algorithms have been proposed and used by physicians in the UK; e.g. the Exeter biologic clinic recommendations for biologic dose reduction (Appendix 5) and recommendations by the NHS Greater Glasgow and Clyde on biologic drug monitoring (Appendix 6). However, to our knowledge, there is no unified treatment algorithm based on TNF testing. Importantly, in the INGEBIO study (conducted in Spain), clinicians were not expected to follow any test-based strategy when making treatment decisions based on test results and clinical judgement. Therefore, it is unclear whether and to what extent the economic results based on this study are relevant to clinical practice in England.
- To our knowledge, there is no standard UK recommendation on managing flares in people with RA.
- The short (18 month) time horizon of the cost-effectiveness analyses undertaken in this study was defined by the observational period in the INGEBIO trial. Given that regular treatment monitoring is a long-term intervention this is a key limitation. Cost and health outcomes were not extrapolated into the future, due to the lack of long-term clinical studies.
- Due to limited reporting, it is not clear to what extent selection bias in the INGEBIO study (which was a non-randomised trial) could have influenced the results of the economic analysis.
- In this study, as in many other economic evaluations in RA, health state utility values were estimated from HAQ scores using published regression functions. It is recognised, however, that the HAQ is a functional measure, and does not capture the

full impact of RA on quality of life. There is also uncertainty in the regression functions used, which has not been explored in this analysis.

- Utility values derived from HRQoL data for rheumatoid arthritis patients were estimated from clinical outcomes in the INGEBIO study which had a *mixed* population of people with RA, psoriatic arthritis (PsA) or ankylosing spondylitis (AS). Since people with RA are usually older and more likely to be female when compared to people with PsA and AS, the utility values for people with RA, used in the economic analysis, are likely to be lower than those for the mixed population (since men tend to value health states higher than women, and the same applies to younger versus older people.[18] This may have overestimated the incremental cost-effectiveness ratios (ICERs).
- Since the rates of AEs were not reported in the INGEBIO study, the impact of AEs was modelled using evidence from another study. However, based on clinical advice and published literature, those AEs which carry a significant cost and disutility burden are relatively rare.
- There was limited evidence on utilities, based on EQ-5D scores, directly relevant to people with RA experiencing flares, people with RA experiencing serious adverse events as well as people with remission, LDA or active disease health status in the UK settings was identified in this study.
- Finally, since the actual costs to the NHS of ADL (Humira®), its biosimilars and other TNF treatments were not known to the AG at the time of writing, the effect of variation in the annual acquisition costs of the biologics within the range of £1,000 - £9,200 per patient was examined in the threshold analysis. However, given that (1) the actual costs of the originator products and their biosimilars vary considerably across England, (2) there is a variation in the uptake of biosimilars across the UK, and (3) the proportion of people treated with biosimilars is likely to increase in the near future due to very recent changes in the biologics market, it is not clear which estimates obtained in our economic analyses are most relevant to the NHS.

Generalisability of the findings

Clinical effectiveness

Given that the best quality study selected in the clinical effectiveness systematic review was conducted in Spain, the generalisability of their findings to the UK NHS setting remains uncertain due to variations in clinical practice and health policies between different countries. Furthermore, the applicability of the findings are limited further because this study, and the

results of changes in therapeutic dose from the historically controlled study by Pascual-Salcedo (2013), were for a mixed population (including RA, PsA and/or AS).

Cost effectiveness

Outcomes from the INGEBIO study were also used in the economic analysis for people in remission/LDA. It was a pragmatic trial, and therefore it is likely that the results could be generalisable to routine practice settings. However, the generalisability to UK clinical practice settings of the findings in the INGEBIO study (in Spain) and therefore the economic results remain uncertain.

Since findings from the mixed population considered in the INGEBIO study might not be generalisable to the RA population, and the quality of this trial was judged to be at serious risk of bias, the economic results presented here should be considered with caution.

Due to the severe paucity of robust studies, no all-test-hits-and-TNF inhibitors from the NICE scope could be modelled using reported clinical outcomes considered in this study. It is therefore not clear whether and to what extent the economic estimates obtained for people treated with ADL are applicable to people treated receiving the other anti-TNF treatments.

According to NHS England documentation, some originator manufacturers have offered discounts, changing the potential for cost savings for the NHS. Therefore, the list prices of TNF inhibitors assumed in the analyses reporting ICERs might not adequately reflect the actual costs of the biologics to the NHS in the coming years.

Conclusions

The findings from this assessment demonstrate very limited evidence on the effect of TDM based on ELISA tests for optimising anti-TNF therapies in people with RA, either in those who had achieved remission or LDA, or in those who had experienced a primary non-response or a secondary non-response.

In relation to clinical effectiveness, limited data were identified evaluating TDM in the target populations. One non-randomised trial compared TDM with standard care (the INGEBIO study) had serious limitations in relation to the NICE scope: only one-third of the participants had RA, many of the analyses were not by intention-to-treat, follow-up was for only 18 months, there was no explicit algorithm for guiding clinicians in how the results of testing should change treatment (e.g. tapering), and the study was only reported in three abstracts. In addition, seven observational studies (reported in eight publications) were also identified but were of minimal value in informing whether ELISA test-based monitoring is clinically effective or not.

Despite these substantial weaknesses in the clinical effectiveness evidence base, a simple model was developed to estimate the cost-utility of ELISA test-based monitoring for people with RA taking bDMARDS. The main effectiveness evidence in the model was also from the poorly reported INGEBIO study, heavily supplemented by input parameters from other studies and expert advice. The results of the economic analysis should therefore be viewed as exploratory and highly speculative. For example, although the INGEBIO study only evaluated testing using Promonitor ELISA kits, for those in remission/LDA treated with Humira® (ADL), with further assumptions these results have been used to estimate the threshold testing costs at which TDM would become cost-effective with people taking other TNF inhibitors (and taking either originator products or biosimilars.

In summary, there is limited valid and applicable research evidence, and much uncertainty in relation to key potential drivers of the effectiveness and cost-effectiveness of using ELISA based testing to monitoring treatment with bDMARDs in people with RA, that no firm conclusions can be drawn

Suggested research priorities

Further controlled trials with a large sample size (preferably RCTs) are required to assess the impact of using the different ELISA tests for monitoring anti-TNF therapies in people with RA who had achieved remission or LDA, and in people being treated with the full range of anti-TNF therapies. We have identified one ongoing Norwegian multicentre RCT (the NOR-DRUM Study) that evaluates the effect of TDM in people with RA in remission compared with standard care. This ongoing trial will provide further useful data on the impact of TDM in the target population.

Future RCTs are warranted to assess the clinical effectiveness of using ELISA tests for monitoring anti-TNF therapies in those people who had developed clinical inefficacy (primary or secondary non-response).

There were no studies identified for people with RA treated with CTZ or GLM. Future RCTs are required to assess the clinical effectiveness of using ELISA tests for monitoring such anti-TNF therapies in the target populations.

Our review identified very limited evidence on healthcare resource use and utilities, based on EQ-5D scores, directly relevant to the patient population considered in this assessment. This warrants further research on medium/long term cost and health outcomes in people with RA treated with the TNF inhibitors.

Plain English Summary

Background

Rheumatoid arthritis (RA) is a chronic systemic autoimmune disease, primarily causing inflammation, pain and stiffness (synovitis) in the joints. Those with severe disease may be treated with biological disease modifying anti-rheumatic drugs (bDMARDs), including the TNF-α inhibitors adalimumab (ADL), etanercept (ETN), infliximab (IFX), certolizumab pegol (CTZ) and golimumab (GLM). Monitoring of response to treatment with TNF-α typically involves clinical assessment and the use of response criteria (DAS28 or EULAR).

Commercial enzyme-linked immunosorbent assay (ELISA) tests can be used to detect and measure drug concentrations and drug antibody levels in the blood. It is thought that these tests have the potential to inform whether adjustments to treatment are required, and also to help clinicians to better understand the reasons for non-response or a loss of response to treatment.

Objective

To evaluate the clinical effectiveness and cost effectiveness of using ELISA tests (Promonitor, IDKmonitor, LISA-TRACKER, RIDASCREEN, MabTrack, and ELISA tests used by Sanquin Diagnostic Services) to measure drug levels and anti-drug antibodies for monitoring response to TNF-α inhibitors (ADL, ETN, IFX, CTZ, GLM) in people with RA who had achieved treatment target (remission or low disease activity [LDA]), or experienced a primary non-response or a secondary non-response.

Methods

Bibliographic literature searching was conducted to identify a published literature reporting clinical outcomes and associated costs of TNF testing. Studies were selected for inclusion versus pre-specified eligibility criteria. An economic analysis was conducted to estimate the short-term health and economic outcomes of adding TNF testing to usual practice to guide treatment decisions in people with RA.

Results

Eight studies (reported in 11 publications) were identified: seven studies investigated dose tapering in patients treated with ADL, ETN, IFX, and one study reported the clinical outcomes of the increase in IFX dose in people who did not respond to treatment. The designs of the identified studies varied greatly, and only one study compared treatment

guided by test results with usual care (i.e. clinical judgments and monitoring using a composite score such as DAS28).

The economic analyses conducted are inconclusive and suggest considerable uncertainty in the cost-effetciveness of therapeutic monitoring of TNF-alpha inhibitors in RA. Data from two reports of the same study produced very different conclusions on the cost-effectiveness of Promonitor testing in people receiving ADL who are in remission/LDA. The results based on the longer follow-up (Arango and colleagues 2017) suggested that monitoring is more costly and produces fewer QALYs than standard care. Of the sensitivity analyses conducted, only one assumption impacted on the results: that the rate of flares alone changes as a consequence of monitoring. This was when evidence from Ucar and colleagues (2017)[11] was used and resulted in standard care dominating the intervention. Exploratory analyses using Promonitor to monitor people in remission/LDA receiving ETN or INF were undertaken, and showed the same results as that for ADL.

Discussion and conclusion

The findings from this assessment demonstrate very limited evidence on the effect of TDM based on ELISA tests for optimising anti-TNF therapies in people with RA, either in those who had achieved remission or LDA, or in those who had experienced a primary non-response or a secondary non-response.

Despite substantial weaknesses in the clinical effectiveness evidence base in the target population, a simple model was developed to estimate the cost-utility of ELISA test-based monitoring for people with RA taking bDMARDS. The main effectiveness evidence in the model was also from the poorly reported INGEBIO study, heavily supplemented by input parameters from other studies and expert advice.

Given substantial uncertainty, the results presented in this study should be considered with caution.

Contents

Abstract 4
Background 4
Methods 4
Results 4
Conclusions 5
Funding 5
Scientific summary 6
Background 6
Objectives 6
Methods 6
Results 9
Discussion 33
Conclusions 37
Plain English Summary 39
Background 39
Objective 39
Methods 39
Results 39
Discussion and conclusion 40
Contents 41
List of tables 45
List of figures 48
Abbreviations 49
Glossary 53
1 Background and definition of the decision problem(s) 54
1.1 Description of the health problem 54
1.1.1 Aetiology, pathology and prognosis 54
1.1.2 Diagnosis of rheumatoid arthritis 55
1.1.3 Epidemiology 56
1.1.4 Impact of health problem 56
1.2 Management of rheumatoid arthritis 58
1.2.1 Treatment of rheumatoid arthritis 58
1.2.2 Monitoring in rheumatoid arthritis 61
1.3 Description of technologies under assessment 63

1.3.1 Summary of technologies 64
1.3.2 Place of tests in the clinical pathway 69
1.3.3 Identification of important sub-groups 69
1.3.4 Current usage in the NHS 69
1.3.5 Anticipated costs associated with the use of the tests 69
1.4 Comparators 70
1.5 Outcomes 70
1.6 Summary of the scope of work 71
2 Assessment of clinical effectiveness 75
2.1 Methods for reviewing effectiveness 75
2.2 Identification of studies 75
2.2.1 Inclusion and exclusion criteria 76
2.2.2 Study selection strategy 80
2.2.3 Data extraction strategy 80
2.2.4 Critical appraisal strategy 81
2.2.5 Methods of data synthesis 81
2.3 Clinical effectiveness results 81
2.3.1 Quantity and quality of research available 81
2.3.2 Assessment of clinical effectiveness 82
2.3.3 Results of clinical effectiveness 100
2.3.4 Discussion 113
3 Systematic review of existing cost-effectiveness evidence 117
3.1 Objectives 117
3.2 Methods 117
3.2.1 Identification of studies 117
3.2.2 Eligibility criteria 118
3.2.3 Data extraction 118
3.2.4 Critical appraisal 118
3.3 Results 118
3.3.1 Non-model based studies 125
3.3.2 Model-based studies 126
3.4 Quality of identified cost-utility studies 134
3.5 Discussion 136
3.6 Conclusions 139
4 Independent economic assessment 140
4.1 Methods 140

4.1.1 Summary of available evidence 140
4.1.2 Choice of modelling approach 143
4.1.3 Analyses 144
4.1.4 Model structure 150
4.1.5 Population 152
4.1.6 Interventions and comparators 153
4.1.7 Perspective, time horizon and discounting 153
4.1.8 Considerations in the development of the independent economic assessment 153
4.1.9 Model parameters 160
4.1.10 Checking the model for wiring errors 197
4.2 Cost effectiveness results 197
4.2.1 Adalimumab and Promonitor 197
4.2.2 Etanercept and Infliximab and Promonitor 208
4.3 Discussion 209
5 Discussion 210
5.1 Statement of principal findings 210
5.1.1 Clinical effectiveness 210
5.1.2 Cost effectiveness 214
5.2 Strengths and limitations of the assessment 215
5.2.1 Clinical effectiveness 215
5.2.2 Cost effectiveness 215
5.3 Uncertainties 218
5.3.1 Clinical effectiveness 218
5.3.2 Cost effectiveness 219
5.4 Generalisability of the findings 220
5.4.1 Clinical effectiveness 220
5.4.2 Cost effectiveness 220
6 Conclusions 222
6.1 Implications for service provision 222
6.2 Suggested research priorities 222
7 References 224
Appendix 1. Literature search strategies 233
Appendix 2. Included and excluded studies 277
Appendix 3. Quality Assessment 330
A3.1 (Part I): Quality assessment on the basis of specific outcomes 330

A3.2 (Part II): Quality assessment of individual studies 332
Appendix 4. Search strategy for the additional search..... 415
Appendix 5. Exeter biologic clinic recommendations for biologic dose reduction..... 418
Appendix 6. Recommendations by NHS Greater Glasgow and Clyde..... 419
Appendix 7. NICE reference case..... 421
Appendix 8. Average cost of joint replacement surgery (the Royal Devon & Exeter NHS Foundation Trust)..... 423
Appendix 9. Estimation of the costs of managing disease health states..... 425

List of tables

Table 1: Clinical effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope 10
Table 2: Cost-effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope 14
Table 3: Model assumptions 21
Table 4: Threshold value for the cost of testing 24
Table 5: Cost-effectiveness results in patients in remission/LDA treated with Humira® and tested using Promonitor 26
Table 6: Additional sensitivity analyses (people in remission/low disease activity) 28
Table 7: One-way deterministic sensitivity analyses 31
Table 8: Cost-effectiveness results for the other tests and TNF inhibitors: people in remission/LDA 32
Table 9: ACR/EULAR 2010 RA classification criteria 55
Table 10: TNF-α inhibitors recommended by NICE for the treatment of severe RA 60
Table 11: Definition of the EULAR response criteria using the DAS28 score 62
Table 12: Test kits under assessment 66
Table 13: Summary of clinical scenarios, drugs, and ELISA technologies 72
Table 14: Characteristics of the included studies – The INGEBO non-randomised controlled study 86
Table 15: Characteristics of the included studies – Observational studies 87
Table 16: Baseline characteristics of included studies – The INGEBO* non-randomised controlled studies 90
Table 17: Baseline characteristics – observational studies 91
Table 18: Characteristics of the Norwegian Drug Monitoring Study (NOR-DRUM)[65] 93
Table 19: Risk of bias in included studies – The INGEBIO non-randomised controlled study 97
Table 20: Risk of bias in included studies – observational studies 97
Table 21: Attrition in the INGEBIO non-randomised controlled study 98
Table 22: Attrition in observational studies 99
Table 23: Changes in disease response, relapse and remission 102
Table 24: Health-related quality of life outcomes 103
Table 25: Changes in disease response, relapse and remission 105
Table 26: Change in disease activity 107
Table 27: Changes in number, direction and magnitude of dose 110
Table 28: Discontinuation of ineffective therapy 112
Table 29: Adverse effects of treatment 112
Table 30: Characteristics of included studies 120
Table 31: Observational cost-effectiveness studies of therapeutic drug monitoring tests in people with rheumatoid arthritis 121
Table 32: Modelling studies of using therapeutic drug monitoring in people with rheumatoid arthritis 123
Table 33: Clinical outcomes from Bartelds and colleagues (2011) 126
Table 34: Cost-effectiveness results reported in Krickaert and colleagues (2015) 129
Table 35: Strategies compared in the Gavan (2017) study 131
Table 36: Quality appraisal of cost-utility studies using the CHEC checklist 135

Table 37: Quality appraisal of cost-utility studies using the checklist developed by Philips and colleagues 136
Table 38: Cost-effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope 137
Table 39: Clinical-effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope 140
Table 40: Clinical outcomes and follow-up period from Ucar and colleagues (2017) and Arango and colleagues (2017) 141
Table 41: Inclusion criteria 142
Table 42: Model assumptions in the analyses with people in remission/low disease activity 146
Table 43: Patient baseline characteristics 152
Table 44: The definition of flares from Maas and colleagues (2013) and Markusse and colleagues (2015) 154
Table 45: Flare duration in the BRASS study 155
Table 46: Mean time to first flare (days) 156
Table 47: Serious adverse events in RA patients treated with anti-TNF inhibitors 158
Table 48: Comparison of effect estimates from consistency and inconsistency models from Singh and colleagues (2015) 160
Table 49: Hospital and Community Health Services combined pay and prices inflation indices 161
Table 50: Acquisition costs of biologics 163
Table 51: Administration cost of infliximab 168
Table 52: HRG codes from the National Schedule of Reference Costs – Year 2017-18, day case 168
Table 53: Resource use and costs of implementing drug level and immunogenicity testing per patient in a hospital setting (Jani and colleagues, 2016) 169
Table 54: Costs of reflex testing provided by the manufacturers of the test kits 172
Table 55: Costs of concurrent testing provided by the manufacturers of the test kits 173
Table 56: Costs of *singlet* and *duplicate* reflex and concurrent testing using Promonitor test kits 176
Table 57: The average six-month resource utilisation stratified by HAQ status (the estimates were based on an average for five years follow-up) 178
Table 58: Unit costs 179
Table 59: HAQ scores for the states of disease activity according to the SDAI, CDAI and DAS28 179
Table 60: Cost of managing disease (Barnabe and colleagues [2013] and Beresniak and colleagues [2011]) 182
Table 61: Average annual costs of managing remission and LDA/active disease based on Barnabe and colleagues (2013) and Beresniak and colleagues (2011) (in GBP 2017-18 prices) 183
Table 62: Cost of managing flares reported in Maravic and colleagues (2005) (2001 prices) 184
Table 63: Definition of disease activity types in Bykerk and colleagues (2014) 187
Table 64: The EULAR response criteria using the DAS28 188
Table 65: Cut-off points to separate the states of remission, and low, moderate and high disease activity using composite indices SDAI and CDAI and DAS28 score 188

Table 66: EQ-5D scores for the states of disease activity according to the SDAI, CDAI and DAS28 189
Table 67: Utility values based on BSRBR data..... 191
Table 68: The definition of flares and their Health Assessment Questionnaire values 192
Table 69: The definition of flares and their utility values 193
Table 70: Threshold value for the cost of testing at which NMB is zero 198
Table 71: Cost-effectiveness results in patients in remission/LDA treated with Humira® and tested using Promonitor 200
Table 72: Health state utility values 203
Table 73: Sensitivity analyses (people in remission/low disease activity) 204
Table 74: One-way deterministic sensitivity analyses based on data from Arango and colleagues (2017) 207
Table 75: Cost-effectiveness results for the other tests and TNF inhibitors: people in remission/LDA..... 208
Table 76: Studies included in the clinical-effectiveness systematic review 277
Table 77: Excluded studies (with reasons) 279
Table 78: Exeter biologic clinic recommendations on dose reduction strategies..... 418
Table 79: Interpretation 419
Table 80: Summary of the reference case..... 421
Table 81: Estimates related to the cost of surgery in people with rheumatoid arthritis (the Royal Devon & Exeter NHS Foundation Trust, April 2017 - September 2018) 423

List of figures

Figure 1: Results of the threshold analyses Arango and colleagues (2017) and Ucar and colleagues (2017) 25
Figure 2: Flow diagram of study inclusion process for the clinical-effectiveness review..... 82
Figure 3: PRISMA flow diagram: a description of study inclusion process for the cost-effectiveness systematic review 119
Figure 4: Algorithm for decision following the result of test in 28th week after adalimumab in Krieckaert and colleagues (2015)..... 127
Figure 5: Algorithm for interpretation of test results 130
Figure 6: Service pathway of RA treatment in England 132
Figure 7: Algorithm for test interpretation used in Gavan (2017) 133
Figure 8: Acquisition costs and QALY change in tapered and non-tapered patients due to flare..... 151
Figure 9: Kaplan-Meier estimates from INGEBIO study 156
Figure 10: Mean HAQ by EULAR response category for patients receiving biologics..... 190
Figure 11: Three types of flares considered in Markusse and colleagues (2015) 192
Figure 12: EQ-5D utility according to British (UK) and Swedish (SD) preference sets of patients with established rheumatoid arthritis treated with TNF-inhibitors..... 194
Figure 13: EQ-5D at first initiation of treatment with TNF-inhibitors for rheumatoid arthritis, psoriatic arthritis, and spondyloarthritis during 2002-08..... 195
Figure 14: EQ-5D during the first courseof treatment with TNF-inhibitors in 2002-08 for people with rheumatoid arthritis, psoriatic arthritis, and spondyloarthritis 195
Figure 15: EQ-5D during follow-up upon withdrawal from treatment..... 196
Figure 16: EQ-5D for people with rheumatoid arthritis (all participants versus participants with complete data) during the first and second lines of anti-TNF-α treatment 197
Figure 17: Results of the threshold analyses using Ucar and colleagues (2017) and Arango and colleagues (2017)..... 199
Figure 18: Interpretation: 3-6/12 after initiation of therapy to guide drug dose/infusion time interval 419
Figure 19: Interpretation: anti-TNF failure of response 420
Figure 20: Distributions and costs for different HAQ bands 425

Abbreviations

Ab	Antibody
ACPA	Anti-citrullinated protein (anti-CCP) antibodies
ACR	American College of Rheumatology
ADAb	Adalimumab antibody
ADL	Adalimumab
AE	Adverse event
AG	Assessment Group
anti-CCP	Anti-cyclic citrullinated peptide
AS	ankylosing spondylitis
AUC	Area under the curve
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BDM	Biologic drug monitoring
bDMARD	Biologic DMARD
BNF	British National Formulary
BRASS	Brigham RA Sequential Study
BSRBR-RA	British Society for Rheumatology Biologics Register – Rheumatoid Arthritis
CDAI	Clinical Disease Activity Index
cDMARD	Conventional DMARD
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CG	Control group
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	Confidence interval
CrI	Credible interval
CRP	c-reactive protein
CTZ	Certolizumab pegol
DAS	Disease Activity Score
DAS28	Disease Activity Score 28 joints
DES	Discrete-event simulation
DL	Drug level
DMARD	Disease-modifying anti-rheumatic drugs

ELISA	Enzyme-linked immunosorbent assay
EQ-5D	European Quality of Life-5 Dimensions
ERAS	Early Rheumatoid Arthritis Study
ESR	Erythrocyte sedimentation rate
ESR	Erythrocyte sedimentation rate
ETN	Etanercept
EULAR	European League Against Rheumatism
FAD	Final appraisal determination
FCE	Finished consultant episodes
GBP	Great Britain Pounds
GC	Glucocorticoid
GLM	Golimumab
HAD	High disease activity
HAQ	Health Assessment Questionnaire
HAQ-DI	Health assessment questionnaire disability index
HL	High level
HR	Hazard ratio
HRG	Healthcare resource groups
HRQoL	Health-related quality of life
HSUV	Health state utility value
HTA	Health Technology Assessment
HUD	Health Utilities Database
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
i.v.	Intravenous
ICER	Incremental cost effectiveness ratio
IFX	Infliximab
IG	Intervention group
IMF	International Monetary Fund
IQR	Interquartile range
IRR	Incidence rate ratio
ITT	Intention to treat
LDA	Low disase activity
LRTI	Lower respiratory tract infectections

MCP	metacarpophalangeal joint
MDA	Moderate disease activity
MTX	Methotrexate
NA	Not applicable
NBT	Non-biologic therapy
NDB	National Data Bank for Rheumatic Diseases
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NMB	Zero net monetary benefit
NOAR	Nortfolk Arthritis Register
NR	Not Reported
NRAS	National Rheumatoid Arthritis Society
NSAID	Non-steroidal anti-inflammatory drug
OR	Odds ratio
PIP	proximal interphalangeal joints
PJP	Pneumocystis jirovecii pneumonia
PPP	Purchasing power parity
PsA	Psoriatic arthritis
QALY	Quality adjusted life years
RA	Rheumatoid Arthritis
RCT	Randomised controlled trial
RF	Rheumatoid factor
RTX	Rituximab
s.c.	Subcutaneous
ScHARR	School of Health and Related Research
SD	Standard deviation
SDAI	Simplified Disease Activity Index
SJCs	Swollen joint counts
SpA	Spondyloarthritis
SSATG	Swedish Arthritis Treatment Group
TA	Technology Appraisal
TDM	Therapeutic drug monitoring
THR	Total hip replacement

TJCs	Tender joint counts
TKR	Total knee replacement
TNF	Tumour necrosis factor
TNF-α	Tumour necrosis factor-alpha
VAS	Visual analogue scale
vs	Versus
WTP	Willingness to pay

Glossary

Assay range	The lowest and highest values within which an assay can detect and quantify the target entity. There will be evidence of acceptable reliability and validity of the test within this range.
Bioequivalence	Where two (or more) drugs have identical active ingredients, similar bioavailability, equivalent physiologic activity and thus interchangeability. Biosimilar drugs demonstrate bioequivalence to an originator product. See Bioequivalence (WHO)
Biosimilar	A biological medicine which is highly similar to another biological medicine already licensed for use. It is a biological medicine which has been shown not to have any clinically meaningful differences from the originator biological medicine in terms of quality, safety and efficacy (Source: NHS England publication, *What is a biosimilar medicine*?)
Brand name	Name given to a pharmaceutical product by the manufacturer, e.g. Valium is the originator brand name (also called trade name) for diazepam. The use of this name is reserved exclusively to its owner as opposed to the generic name, i.e. diazepam. Brand names may also be used for generic products; they are then often called 'branded generics'. These brand names are different from innovator brand names. See Generic medicine (WHO)

1 Background and definition of the decision problem(s)

1.1 Description of the health problem

Rheumatoid arthritis (RA) is a systemic autoimmune disease, primarily causing chronic inflammation and destruction of the joints. The disease usually has a relapsing-remitting course, involving flare-ups followed by periods of lower disease activity (LDA). However, for some people RA is constantly progressive and for others the disease might be short-lived.[19] Whether or not periods of remission or LDA are achieved, RA requires monitoring, to enable appropriate adjustments to be made to treatment.

1.1.1 Aetiology, pathology and prognosis

RA typically affects the synovial tissue of the small joints of the hands and feet. However, any synovial joint may be affected, causing swelling, stiffness and pain (synovitis), and progressive joint destruction. As a systemic disease, the whole body may be affected, including the lungs, heart and eyes. Systemic symptoms may include a non-specific feeling of general illness, fatigue, systemic inflammation and depression.[20,21]

The underlying reasons for the development of RA are complex and not fully understood. It is clear, however, that both genetic and environmental factors are involved. Genetic factors contribute an estimated two-thirds of the risk of developing RA,[22] and also influence the progression and severity of the disease.[21,22] Non-genetic factors that increase the risk of developing RA include: female sex (potentially due to hormonal factors, with lowered risk of developing the disease during pregnancy, with oral contraceptive use, and in women who have breastfed, although this latter relationship is somewhat less clear);[22-24] regular smoking (this relationship is dose-dependent,[22] male smokers are particularly susceptible,[25] and smokers also experience more severe RA symptoms;[26] dietary factors and obesity (including a high intake of red meat, salt and free fructose, and a low intake of vitamin C containing fruits and vegetables);[23,27] periodontitis;[22] and advanced age.[22]

For people with RA, these complex genetic and environmental factors lead to repeated activation of the innate and adaptive immune systems, leading to poor immune self-tolerance, the activation of antigen-specific T and B cells, and the production of antibodies associated with RA (rheumatoid factor (RF) and anti-cyclic citrullinated peptide (anti-CCP)). These changes contribute to the destruction of the synovial joints and the other inflammatory symptoms seen in RA.[21] It is now known that dysregulation in the production of tumour necrosis factor-alpha (TNF-α) (a cell signalling protein that promotes inflammatory responses), can contribute to inflammatory disease; TNF-α is implicated in the development of many of the symptoms of RA (joint pain and destruction, fatigue, and weight loss).

There is no cure for RA and there is substantial individual variation in the course of the disease. RA may be short-lived (i.e. achieving remission with no evidence of disease), relapsing-remitting (patterns of flare-ups followed by periods of improvement) or may be refractory despite treatment, with disease continually worsening.[19] Data published in 2004 suggest that, whilst 10-15% of people with RA have refractory RA, and 10-15% experience full remission within five years of treatment, 70-80% have relapsing-remitting disease,[28] Newer data suggests that remission rates are increasing and symptom flare-ups decreasing, principally in the first five years after diagnosis. However, the majority of people with RA are still experiencing relapsing-remitting disease.[29]

1.1.2 Diagnosis of rheumatoid arthritis

A diagnosis of RA usually involves both laboratory tests and an assessment of clinical signs and symptoms. According to National Institute for Health and Care Excellence (NICE) guidance on the management of RA in adults (NG100), initial testing should include blood tests for rheumatoid factor (RF) and x-rays of the hands and feet. Additionally, C-reactive protein (CRP) testing should be considered for those with negative RF results.[30]

To aid clinical diagnosis, the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) have developed classification criteria for RA. These criteria attribute points based on the number of tender or swollen joints, serologic tests for RF and anti-CCP antibodies and tests for acute phase reactants (CRP and erythrocyte sedimentation rate (ESR)). The duration of symptoms is also assessed (Table 9). A total score of ≥6 points (currently or previously) on the ACR/EULAR classification system, together with clinically obvious synovitis, is considered to indicate definite RA if symptoms cannot be better explained by an alternative diagnosis.[31]

Table 9: ACR/EULAR 2010 RA classification criteria

Joint distribution	**Points (0-5)**
1 Large joint	0
2–10 Large joints	1
1–3 Small joints (large joints not counted)	2
4–10 Small joints (large joints not counted)	3
>10 Joints (at least one small joint)	5
Serology	**Points (0–3)**
Negative RF *and* negative ACPA	0
Low positive RF *or* low positive ACPA	2
High positive RF *or* high positive ACPA	3
Symptom duration	**Points (0-1)**
<6 weeks	0
>6 weeks	1

Acute phase reactants	Points (0-1)
Normal CRP *and* normal ESR	0
Abnormal CRP *or* abnormal ESR	1

Key: RF: rheumatoid factor; ACPA: anti-citrullinated protein (anti-CCP) antibodies; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate.

1.1.3 Epidemiology

Estimates made in 2002, indicate that there are 400,000 people in England and Wales living with RA, with 10,000 incident cases per year.[32,33] However, the figure is likely higher, with data from 2009 suggesting that, in England alone, RA affects approximately 0.8% of the population, or 580,000 adults, with 26,000 new cases diagnosed each year.[19] RA is approximately three times more prevalent in women than in men (see Section 1.1.1), has a peak age of incidence in the UK of 70-79 years,[30] and is less prevalent amongst people with a higher educational level and people in non-manual employment.[22]

According to data from The British Society for Rheumatology Biologics Register for RA (BSRBR-RA), between 2001 and 2014, 13,502 people with RA began treatment with a TNF-α inhibitor, although the number will be higher as not all people treated with biologics are recruited to the BSRBR-RA study and not everyone consents to inclusion.[34] Consistent with RA as a whole, 76% were female. Median age of those starting TNF-α inhibitor therapy was 57 years (IQR 49–65) and disease was severe, with a median disease activity score in 28 joints (DAS28) score of 6.5 (IQR 5.8–7.2) (see Section 1.2.2.1.1).[34]

Historically, there has been concern about geographical variation in access to TNF-α inhibitors. Although available data on this are not up to date, and despite geographical variation in service provision, differential geographical access to biologic treatment for RA is no longer considered an issue. There is, however, recent evidence to suggest that choice of specific TNF-α inhibitor in England might be influenced by age and relationship status.[35]

1.1.4 Impact of health problem

RA varies greatly from person to person, but often results in substantial morbidity, impaired physical activity, and poor quality of life, leading to a reduced life expectancy (although increased mortality has been decreasing over time).[36]

The disease is often multi-morbid; data published in 2006 from the BSRBR-RA suggests that, amongst people treated with biological agents, 58% have at least one comorbid condition, most commonly hypertension, depression, peptic ulcer disease, and respiratory disease.[37] Due to the chronic nature of RA, coupled with the high-risk of co-morbidities,[37] a multidisciplinary team of health professionals and services are required for the management

of the disease.[30] Support may also be sought from patient groups. RA is, therefore, associated with a substantial cost burden to the NHS. A report by the National Rheumatoid Arthritis Society (NRAS), published in 2010, estimates the annual cost to the NHS of RA, including the costs of drug acquisition and hospitalisation, to be nearly £700 million.[38]

With regard to indirect costs, approximately one-third of people with RA stop work within two years of the onset of symptoms, and this prevalence increases thereafter. For those remaining in work, sickness absence is greater amongst people with RA compared with people without RA (40 days versus 6.5 days per year).[19] The 2010 NRAS report estimates that annual productivity losses due to RA in England and Wales totals over £7 billion.[38]

Based on costs in the British National Formulary (BNF) 2018, the costs to the NHS of TNF-α inhibitors per patient per year are:

- £9,187.08 for adalimumab (ADL) (Humira®)
- £9,155.64 for golimumab (GLM) (Simponi®)
- £9,326.92 for certolizumab pegol (CTZ) (Cimzia®; although the cost in the first year is £10,399.42)
- £9,326.92 for etanercept (ETN) (Enbrel®)
- £8,557.29 and £8,394.23 for ETN biosimliars Benepali® and Erelzi® respectively
- £5,747.48 for infliximab (IFX) (Remicade®; £7,730.18 in the first year)
- £5,172.76 for IFX biosimilars Inflectra® and Remsina® (£6,957.20 in the first year) and
- £5,163.72 for IFX biosimilar Flixabi® (£6,945.05 in the first year)

Costs will vary with dosing changes or due to negotiated procurement discounts. It should be noted that the cost of ADL has very recently decreased, due to the approval of biosimilars (Amgevita®, Hulio®, Imraldi® and Hyrimoz®), although these costs could not be accessed and estimated percentage uptake of these products was unclear at the time of writing. There is also a substantive wastage cost associated with biologic treatments, averaging an estimated £370 per patient per year.[39] When people continue to be prescribed TNF-α inhibitors unnecessarily, there is an obvious cost implication. Unnecessary continued treatment may also lead to unnecessary side-effects. Potential side effects of TNF-α inhibitors may include, but are not limited to, increased risk of viral and bacterial infections (of the respiratory tract, bladder and skin), allergic reactions, nausea and vomiting, itching, and fever (see Table 10 for very common adverse reactions). Efficient systems for monitoring response to these

treatments, and thus informing decisions on optimal drug dosing or on treatment discontinuation, could therefore be of benefit to the NHS.

1.2 Management of rheumatoid arthritis

According to the NICE guidance for RA (2018), and the NICE RA care pathway,[30,40] active RA in adults should be treated with the aim of achieving a target of remission or low disease activity (LDA) (treat-to-target).The main aim of treatment and management of RA is, therefore, to achieve target symptom control and to prevent further damage. Monitoring of treatment response is required to enable appropriate treatment adjustments to be made.

1.2.1 Treatment of rheumatoid arthritis

The NICE guidance for RA recommends the use of disease modifying anti-rheumatic drugs (DMARDs).[30] Short-term (bridging) glucocorticoids might be offered prior to starting DMARDs. Where control of pain and inflammation is inadequate, non-steroidal anti-inflammatory drugs (NSAIDS, including cox II selective inhibitors) are used, sometimes in combination with other analgesics. In established disease, complications and associated comorbidities are addressed and treated as appropriate. This may involve physiotherapy, occupational therapy, podiatry, psychological therapies, complementary therapies and dietary advice, and in persistent and worsening cases of joint damage, surgery may be offered.[30]

Disease modifying treatment may be broadly classified as conventional (cDMARDs; including methotrexate, leflunomide, sulfasalazine, and hydroxychloroquine), synthetic (sDMARDs; such as the Janus kinase inhibitor tofacitinib) or biologic (bDMARDs; including, but not limited to, TNF-α inhibitors). The NICE guidance for RA (2018) and the NICE RA care pathway indicate that initial DMARD treatment for adults with active RA should begin with cDMARD monotherapy, if possible within three months of symptom onset. If treatment targets are not met, despite dose escalation, further cDMARDs are added.[30,40]

1.2.1.1 The role of TNF-α inhibitors in the care pathway

The NICE care pathway states that bDMARDs, including TNF-α inhibitors, should only be offered to people with severe disease that has not been controlled with cDMARDs.[30,40] NICE Technology Appraisal (TA) 375 guidance[8] recommends ADL, ETN, IFX, CTZ and GLM, in combination with methotrexate, for use in severe RA (i.e. DAS28 > 5.1; see Section 1.2.2.1) that has not responded to intensive therapy with at least two cDMARDs, including methotrexate. ADL, ETN and CTZ may also be used as monotherapy in people for whom methotrexate is contraindicated or not tolerated. As part of TA375, NICE also makes

recommendations for two other bDMARDs (tocilizumab and abatacept),[8] but these interventions are ouside of the scope of this appraisal..

A summary of the recommended TNF-α inhibitors relevant to this report, their contraindications and very common adverse reactions, and a list of biosimilars, is provided in Table 10. The biosimilars listed in Table 10 are thought to have bioequivalence (and are also often assumed to perform similarly) to the reference/originator products.[41] It should be noted that IFX is administered by an intravenous infusion in the outpatient setting, whereas the other recommended TNF-α inhibitors may be self-administered by subcutaneous injection (usually administered by patients in their own homes). TA375 recommends that treatment should start with the least expensive drug (taking into account administration costs, dose needed and product price per dose).[8]

Table 10: TNF-α inhibitors recommended by NICE for the treatment of severe RA

TNF-α inhibitors	Recommended use[1]	Contraindications	Very common adverse reactions	Administration	Brand names[2]
ETN	In combination with methotrexate, for use in severe RA (i.e. DAS28 > 5.1) - as monotherapy when methotrexate is contraindicated or not tolerated	Sepsis or risk of sepsis, active infections (chronic or localised)	Infections and injection site reactions	Subcutaneous injection; 50 mg weekly or 25 mg twice weekly	Enbrel*, Erelzi, Benepali, Lifmior, Brenzys
ADL	In combination with methotrexate, for use in severe RA (i.e. DAS28 > 5.1)) - as monotherapy when methotrexate is contraindicated or not tolerated	Active tuberculosis, other severe infections, moderate to severe heart failure	Respiratory tract infections, leukopenia, anaemia, increased lipids, headache, abdominal pain, nausea and vomiting, elevated liver enzymes, rash, musculoskeletal pain, injection site reaction	Subcutaneous injection; 40 mg every other week	Humira*, Amgevita, Cyltezo, Imraldi, Solymbic, Hyrimoz, Halimatoz
IFX	In combination with methotrexate, for use in severe RA (i.e. DAS28 > 5.1)	Active tuberculosis, other severe infections, moderate to severe heart failure	Viral infection, headache, upper respiratory tract infection, sinusitis, abdominal pain, nausea, infusion-related reaction and pain	Intravenous Infusion; 3 mg/kg at 0, 2 and 6 weeks, and then every 8 weeks[3]	Remicade*, Inflectra, Remsima, Flixabi, Zessly, Renflexis, Ixifi
CTZ	In combination with methotrexate, for use in severe RA (i.e. DAS28 > 5.1)) - as monotherapy when methotrexate is contraindicated or not tolerated	Active tuberculosis, other severe infections, moderate to severe heart failure	None listed[4]	Subcutaneous injection; 400 mg at 0, 2 and 4 weeks, and then 200 mg every 2 weeks[5]	Cimzia*
GLM	In combination with methotrexate, for use in severe RA (i.e. DAS28 > 5.1)	Active tuberculosis, other severe infections, moderate to severe heart failure	Upper respiratory tract infections	Subcutaneous injection; 50 mg monthly[6]	Simponi*

Key: ADL: adalimumab; CTZ: certolizumab pegol; DAS28: disease activity score in 28 joints; ETN: etanercept; GLM: golimumab; IFX: infliximab
Notes:
[1] Recommended in NICE Technology Appraisal (TA) 375
[2] Brand names include both originator and biosimilar products, the originator/ reference products are denoted with a * whilst the remaining brand names refer to biosimilars.
[3] If there is inadequate response or loss of response after 12 weeks, there may be a stepped increase in dose of 1.5 mg/kg up to 7.5 mg/kg every eight weeks, or an increase of administration of the 3 mg/kg dose to every four weeks.
[4] No very common adverse reactions listed in the summary of product characteristics, but in clinical trials the most common adverse reactions were bacterial and viral infections.
[5] Following confirmed clinical response, a dose of 400 mg every four weeks may be given.
[6] If there is inadequate response after three to four doses, dose may be increased to 100 mg in people weighing over 100 kg.

Whilst TNF-α inhibitors have been found to be of benefit in the treatment of RA,[8] some people do not respond to these treatments (primary non-responders), and others experience a loss of response (secondary non-responders). Secondary non-response may be due to changes in the disease, the development of antibodies to the TNF-α inhibitor, or fluctuations in circulating drug levels.

1.2.2 Monitoring in rheumatoid arthritis

Monitoring in RA can be used to identify primary and secondary non response, potentially improving access to specialist services and informing treatment alteration decisions. Monitoring can also be used to make treatment adjustments for those who have achieved treatment targets. The 2018 NICE guidance for RA recommends a monitoring review appointment six months after treatment targets are achieved, to ensure maintenance of the target. Monitoring should continue annually to assess disease activity, treatment response, functioning, impact on quality of life, comorbidities, complications and the need for surgery, and to arrange multidisciplinary referrals.[30]

1.2.2.1 Current methods for monitoring treatment response

Due to the huge variation between indiciduals in the severity and course of RA, and thus in treatment targets, it is incredibly difficult to measure changes in the disease in a standardised way. Indeed, in clinical practice, evaluation of both treatment response and symptom flare ups is multifaceted, and may involve assessment of a number of domains (pain, fatigue, activity level, overall physical and mental health, functioning in work and education, complications and adverse events [AEs]) in addition to measuring disease activity (using standardised scales and additional imaging).

There are a range of classification systems and scales that have been developed to measure and monitor disease activity in RA (as well as scales that are commonly used to measure other domains such as disability or activity level, such as the Health Assessment Questionnaire (HAQ).[42] Disease activity is commonly measured using: clinical examination, such as swollen joint counts (SJCs) and tender joint counts (TJCs); laboratory test results (e.g. CRP or ESR); or composite measures based on a combination of the above, such as DAS28,[43] the Clinical Disease Activity Index (CDAI),[44] the Simplified Disease Activity Index (SDAI),[44] the ACR20 improvement criteria,[45] and the EULAR response classification system.[46]

In current clinical practice, the DAS28 scales and the EULAR response classification system (which is based on the DAS28) are most commonly used to monitor disease activity. The

use of ultrasound is not recommended for routine monitoring of disease activity in adults with RA.[30,47]

1.2.2.1.1 DAS28

There are two variations of the DAS28, the DAS28-ESR and the DAS28-CRP.[43] Both scales are composite scores that assess 28 joints (shoulder, knee, elbow, wrist, metacarpophalangeal joint (MCP) joints one to five, proximal interphalangeal joints (PIP) joints one to five, bilaterally) for swelling (SW28) and tenderness to touch (TEN28), and also involve the patient's self- assessment of disease activity in the past week on a scale of 0-100. Both scales additionally include blood markers of inflammation (ESR for the DAS28-ESR and CRP for the DAS28-CRP).

Overall disease activity scores are calculated as follows:

- DAS28-ESR = 0.56 × $TEN28_{0.5}$ + 28 × $SW28_{0.5}$ + 0.70 × ln (ESR) + 0.014 × SA.
- DAS28-CRP = [0.56 × sqrt(TEN28) + 0.28 × sqrt(SW28) + 0.36 × ln(CRP + 1)] × 1.10 + 1.15

A DAS28 score >5.1 denotes severe disease activity, ≤ 5.1 but >3.2 moderate disease activity, ≤ 3.2 but ≥2.6 LDA and <2.6 disease remission.[48,49]

1.2.2.1.2 EULAR response classification

The EULAR response classification system is based on improvement in DAS28 scores from initial measurement.[46] The EULAR system classifies improvement as either 'none', 'moderate' or 'good'. The relationship between the DAS28 and the EULAR response classifications are provided in Table 11.

Table 11: Definition of the EULAR response criteria using the DAS28 score

DAS28 at endpoint	Improvement in DAS28 ≤ 1.2	Improvement in DAS28 > 0.6 and ≤ 1.2	Improvement in DAS28 ≤ 0.6
≤3.2	good	moderate	none
>3.2 and ≤5.1	moderate	moderate	none
>5.1	moderate	none	none

Note: This table contains information from Stevenson and colleagues (2016),[50]

1.2.2.2 Monitoring of response to TNF-α inhibitors

Although monitoring of response to treatment with TNF-α inhibitors typically involves the systems described above (clinical assessment, DAS28, EULAR response criteria), there are

neither gold standards nor guidelines available specifically regarding the monitoring of TNF-α inhibitors. More recently, biochemical ELISA testing has emerged to measure blood levels of TNF-α inhibitors, or antibodies to TNF-α inhibitors in people with RA. These testing kits and services (LISA-TRACKER, IDKmonitor, RIDASCREEN, MabTrack and Promonitor kits, and ELISAs used by Sanquin Diagnostic Services) might be useful for detecting primary and secondary non-response to TNF-α inhibitors and in the optimisation of dosages for those who are responding well. For those whose response to therapy has waned, the results of the tests are frequently dichotomised using a cut-off assay result: people may thus be classified as having either therapeutic or sub-therapeutic levels of the drugs, or may be classified as having clinically significant or insignificant levels of antibodies.

These tests may also elucidate reasons for treatment non-response. For example, non-adherence to TNF-α inhibitors may play a part in failure to respond to treatment. Monitoring of blood levels of TNF-α inhibitors, or antibodies to TNF-α inhibitors, can help to reveal non-adherence. In a three-year study assessing non-adherence to ETN (using ELISA testing) in people with RA, 4.1% (95% CI 2.2–7.2) were non-adherent to treatment (non-adherence defined as serum ETN trough concentration <0.1 ug/mL in the absence of a valid medical reason), and 3.4% (95% CI 0.8–10.4) of treatment non-responders had insufficient etanercept exposure, indicative of non-adherence.[51]

The administration of TNF-α inhibitor and anti-drug antibody assays most frequently occurs just before the next administration of the TNF-α inhibitor. This enables simultaneous measurement of a ‘trough’ level of the drug. The tests may be conducted concurrently, or using a reflex testing strategy where the test for TNF-α-inhibitor drug levels is conducted first and the result used to guide follow-up testing by the laboratory without a further request from the treating clinician (i.e. TNF-α inhibitor antibody testing would be only be conducted when the drug was not detected in the sample).

1.3 Description of technologies under assessment

The purpose of this work is to provide NICE with the most up-to-date evidence on the effectiveness and cost-effectiveness of alternative testing and monitoring approaches for assessing TNF-α inhibitor levels and antibodies to TNF-α inhibitor levels, in people with RA undergoing treatment with ADL, ETN, IFX, CTZ, or GLM in the UK. There are three clinical scenarios in which the tests in scope of this appraisal (Section 1.3.1) may be used: (i) remission/LDA to check whether continued treatment at the same dose is appropriate; (ii) primary non-responders (defined as those who have little to no improvement in clinical signs and symptoms initially and as treatment continues), and; (iii) secondary non-responders

(people with an initial response to a TNF-α inhibitor followed by loss of efficacy). Testing could help clinicians and patients to understand the reasons for a non-response or loss of response.

1.3.1 Summary of technologies

The technologies to be evaluated are biochemical ELISA testing kits and services, for measuring levels of TNF-α inhibitors or antibodies to TNF-α inhibitors, typically in the period immediately before administration of the next dose (i.e. trough levels), conducted in addition to current clinical practice in the UK (i.e. clinical assessment and monitoring using a composite score such as DAS28, see Sections 1.2.2.1 and 1.4).

There are six companies providing different test kits or services for up to five TNF-α inhibitors or the antibodies to those TNF-α inhibitors. The test kits are summarised in Table 12. In addition to these test kits, the service provided by Sanquin Diagnostic Services (testing service using validated ELISAs), covering ADL, CTZ, ETN, GLM and IFX drug levels and ETN anti-drug antibodies will be evaluated. Further detail on these test kits and services are provided in Sections 1.3.1.1 to 1.3.1.5. It should be noted that although several of the ELISA tests measure the same drugs (and drug antibodies) there is significant variation between tests in their assay (detection) ranges. This means that some tests may be able to detect and quantify lower and/or higher levels of the same analyte than others.

1.3.1.1 Promonitor (Grifols–Progenika)

Promonitor (Grifols–Progenika) is a portfolio of assays that measure drug levels (ETN, IFX and IFX biosimilars, ADL, GLM) and their correlating anti-drug antibodies (anti-ETN, anti-IFX, anti-ADL, anti-GLM), see Table 12. The kits are manufactured by Proteomika and distributed in the UK by Grifols UK. They consist of strips of pre-coated microtitre plate (96 wells), reagents, buffers, standards, controls and ELISA cover films. The ELISA tests are laboratory-based, conducted either manually or on an automated ELISA processor.

1.3.1.2 IDKmonitor ELISA kits (Immundiagnostik/BioHit Healthcare)

IDKmonitor ELISA kits are manufactured by Immundiagnostik AG and distributed in the UK by Biohit Healthcare Ltd. The ten kits measure either levels of free TNF-α inhibitor or free anti-drug antibodies or total levels of anti-drug antibodies (free antibodies and antibodies bound to the drug), see Table 12. The kits consist of strips of pre-coated microtitre plate (96 wells), reagents, buffers, standards (drug level ELISAs only) and controls. The ELISA tests are laboratory-based, conducted either manually or on an automated ELISA processor.

1.3.1.3 LISA-TRACKER ELISA kits (Theradiag)

LISA-TRACKER ELISA kits are manufactured by Theradiag. The ten kits measure either levels of free anti-drug antibodies or levels of free TNF-α inhibitor, see Table 12. In addition, LISA-TRACKER Duo kits are available (these include assays to measure the levels of both free anti-drug antibodies and the TNF-α inhibitor). The LISA-TRACKER ELISA kits consist of pre-coated strips of microtitre plate (96 wells), reagents, wash buffer, standards and controls. They are laboratory based assays that can be run simultaneously or individually on any manual or automated standard ELISA-based processor platform.

1.3.1.4 RIDASCREEN (R-Biopharm)

The RIDASCREEN enzyme linked immunoassays are manufactured by R-Biopharm. The four kits are laboratory based assays measuring either levels of free TNF-α inhibitor or free anti-drug antibodies, see Table 12. The RIDASCREEN ELISAs are commercialised versions of the KU Leuven in-house ELISAs, and are marketed as apDia ELISA kits in the Benelux area of Europe.

1.3.1.5 MabTrack ELISA kits and Sanquin Diagnostic Services

Sanquin is a laboratory in the Netherlands providing laboratory test services including testing for TNF-α inhibitors using ELISA based assays. The testing service using validated ELISAs is available for etanercept and its correlating anti-drug antibodies, golimumab drug levels and CTZ drug levels. It also provides CE marked MabTrack ELISA kits for local laboratory testing for ADL and IFX levels and their correlating anti-drug antibodies, see Table 12. The MabTrack ELISA kits consist of pre-coated strips of microtitre plate (96 wells), reagents, wash buffer, standards or calibrators, controls and ELISA cover films.

Table 12: Test kits under assessment

Technologies	Company	Variations/kits	Drug/antibodies assessed
Promonitor ELISA kits	Grifols - Progenika	Promonitor-ADL-1DV (50802300DV)	Free[1] ADL
		Promonitor-ANTI-ADL-1DV (50902300DV)	Free[1] anti-ADL antibodies
		Promonitor-ETN-1DV (51102300DV)	Free[1] ETN
		Promonitor-ANTI-ETN-1DV (51202300DV)	Free[1] anti-ETN antibodies
		Promonitor- IFX-1DV (50802300DV)	Free[1] IFX (Remicade®, and biosimilars)
		Promonitor-ANTI-IFX-1DV (50702300DV)	Free[1] anti-IFX antibodies
		Promonitor-GLM-1DV (52002300DV)	Free[1] GLM
		Promonitor-ANTI-GLM-1DV (52102300DV)	Free[1] anti-GLM antibodies
IDKmonitor ELISA kits	Immundiagnostik/BioHit Healthcare	IDKmonitor infliximab drug level ELISA (K9655)	Free[1] IFX (Remicade®, Remsima®, Inflectra®)
		IDKmonitor adalimumab drug level ELISA (K9657)	Free[1] ADL
		IDKmonitor etanercept drug level ELISA (K9646)	Free[1] ETN
		IDKmonitor golimumab drug level ELISA (K9656)	Free[1] GLM
		IDKmonitor infliximab free ADA ELISA ((K9650)	Free[1] anti-infliximab antibodies

Technologies	Company	Variations/kits	Drug/antibodies assessed
		IDKmonitor adalimumab free ADA ELISA (K9652)	Free[1] anti-ADL antibodies
		IDKmonitor etanercept free ADA ELISA (K9653)	Free[1] anti-ETN antibodies
		IDKmonitor golimumab free ADA ELISA (K9649)	Free[1] anti-GLM antibodies
		IDKmonitor infliximab total ADA ELISA (K9654)	Total[2] anti-IFXantibodies
		IDKmonitor adalimumab total ADA ELISA (K9651)	Total[2] anti-ADL antibodies
LISA-TRACKER kits	Theradiag	LISA-TRACKER adalimumab (LTA002)	Free[1] ADL
		LISA-TRACKER certolizumab (LTC 002)	Free[1] CTZ
		LISA-TRACKER etanercept (LTE 002)	Free[1] ETN
		LISA-TRACKER infliximab (LTI002)	Free[1] IFX
		LISA-TRACKER golimumab (LTG002)	Free[1] GLM
		LISA-TRACKER anti-adalimumab (LTA003)	Free[1] anti-ADL antibodies
		LISA-TRACKER anti-certolizumab (LTC003)	Free[1] anti-CTZ antibodies
		LISA-TRACKER anti-infliximab (LTI003)	Free[1] anti-IFX antibodies
		LISA-TRACKER anti-etanercept (LTE003)	Free[1] anti-ETN antibodies

Technologies	Company	Variations/kits	Drug/antibodies assessed
		LISA-TRACKER anti-golimumab (LTG003)	Free[1] anti-GLM antibodies
		LISA-TRACKER Duo adalimumab (LTA005)	Total[2] ADL
		LISA-TRACKER Duo certolizumab (LTC005)	Total[2] CTZ
		LISA-TRACKER Duo etanercept (LTE005)	Total[2] ETN
		LISA-TRACKER Duo Infliximab (LTI005)	Total[2] IFX
RIDASCREEN	R-Biopharm	RIDASCREEN ADM monitoring	Free[1] ADL
		RIDASCREEN anti-ADM antibodies	Free[1] antibodies to ADL
		RIDASCREEN IFX monitoring	Free[1] IFX (Remicade®, Remsima®, Inflectra®)
		RIDASCREEN anti-IFX antibodies	Free[1] antibodies to IFX
MabTrack ELISA kits	Sanquin	MabTrack level adalimumab M2910	Free[1] ADL
		MabTrack ADA adalimumab M2950	Free[1] antibodies to ADL
		MabTrack level infliximab M2920	Free[1] IFX (Remicade®, Remsima®, Inflectra®)
		MabTrack ADA infliximab M2960	Free[1] antibodies to IFX

Key: ADL: adalimumab; CTZ: certolizumab pegol; ETN: etanercept; GLM: golimumab; IFX: infliximab

Notes:

[1] Free TNF-α inhibitor is drug that is unbound to antibody, and free anti-drug antibodies are those that are unbound to drug.

[2] Total anti-drug antibodies include both unbound (free) antibodies and those bound to TNF-α inhibitor.

1.3.2 Place of tests in the clinical pathway

NICE guidance (TA375[8]) states that treatment with a TNF-α inhibitor should only be continued if there is a moderate initial response (using EULAR criteria) at 6 months after treatment initiation and that treatment should be withdrawn if a moderate EULAR response is not maintained.[8] NICE also provides guidance (TA195) on the treatment of RA after a TNF-α inhibitor has failed.[52] The addition of ELISA testing kits to current clinical monitoring procedures has the potential to inform decisions about treatment continuation, treatment optimisation. In addition, ELISA testing may also help cinicians tounderstand the reasons for non-response or loss of response, inform decisions on dosing, and enable adherence to treatment to be assessed. As such, the ELISA testing kits fall into the monitoring and review (following drug treatment) section of the NICE care pathway.[40]

1.3.3 Identification of important sub-groups

People with RA can be grouped according to three clinical scenarios: primary non-response, secondary non-response and remission. However, with regards to particular patient characteristics, there are no subgroups for which the tests are expected to perform differently.

1.3.4 Current usage in the NHS

In UK clinical practice, the tests under assessment (Section 1.3.1) are currently not routinely used for people with RA, and are performed in only two UK laboratories (Viapath and Exeter Clinical Lab). At the Exeter Laboratory, TNF testing is done by using IDKmonitor test kits, while LISA-TRACKER ELISA assays are used at Viapath. However, even these are currently used ad hoc to assist in making treatment management decisions; e.g. dose adjustment rather than being used in routine monitoring strategies.

1.3.5 Anticipated costs associated with the use of the tests

The costs of the ELISA kits and services are detailed in Section 4.1.9.1.8. In addition to the costs of the tests themselves, and based on a recent micro-costing study,[53] the following costs have been identified as being associated with the use of these tests:

- Pre-testing phase: a single outpatient appointment with a consultant rheumatologist and a follow-up appointment with a phlebotomist or clinical support worker
- Analysis phase: costs associated with personnel time, any additional materials required to analyse patient samples (excluding assumed costs such as equipment costs, overhead costs, and capital costs)

- Treatment decision stage: cost of interpretation of test results by a consultant rheumatologist, cost of a telephone discussion of the results with the patient, cost of a letter outlining results and treatment decisions.

These costs are described in further detail in section 4.1.9.1.8.

1.4 Comparators

Comparison will be made between monitoring strategies that use the index tests or services described in Section 1.3.1 (in addition to current clinical practice in the UK) and current clinical practice alone (i.e. clinical assessment and monitoring using a composite score such as DAS28, ACR response criteria or EULAR response criteria).

Currently used monitoring strategies are described in Section 1.2.2.1.

1.5 Outcomes

The outcomes of interest in the assessment of clinical effectiveness included:

i. test (procedural) outcomes: number of inconclusive test results and time to test result;

ii. ii) treatment and management outcomes: number, direction and magnitude of dose changes, frequency of dose adjustments (e.g. dose reduction) due to monitoring, frequency of treatment switching to an alternative biologic, discontinuation of ineffective treatment

iii. clinical outcomes: measures of change in disease activity, rates and duration of disease response, relapse and remission; rates of surgical intervention, rates of hospitalisation, and adverse effects (AEs) of treatment

iv. patient-related outcomes: health related quality of life (HRQoL).

The cost-effectiveness modelling took into account costs/resource use and patient outcomes. The main cost considerations were categorised as costs incurred through the acquisition and administration of biologics, costs associated with testing (drug trough levels and anti-drug antibodies), and the cost of disease management for each health state. The relevant patient outcomes that informed the economic model were the percentage on tapered dose (remission), the rate of flare, and the rate of AEs. The economic modelling considered both concurrent and reflex testing and how the frequency of testing may impact upon cost-effectiveness.

1.6 Summary of the scope of work

In summary, this work evaluated theclinical- and cost-effectiveness of the testing kits and services described in Section 1.3.1, in people with RA undergoing treatment with ADL, ETN, IFX, CTZ, or GLM in the UK. A summary of the clinical scenarios in which each test might be used, and thus the scope of the work, is provided in Table 13.

Table 13: Summary of clinical scenarios, drugs, and ELISA technologies

Clinical scenario	TNF-α inhibitor	Drug/Antibody	ELISAs					
			Promonitor	IDKmonitor	LISA-TRACKER	RIDASCREEN	MabTrack	Sanquin
Remission	ADL	Drug	X	X	X	X	X	X
		Antibody	X	X	X	X	X	
	ETN	Drug	X	X	X			X
		Antibody	X	X	X			X
	IFX	Drug	X	X	X	X	X	X
		Antibody	X	X	X	X	X	
	GLM	Drug	X	X	X			X
		Antibody	X	X	X			
	CTZ	Drug			X			X
		Antibody			X			
Primary non-responder	ADL	Drug	X	X	X	X	X	X
		Antibody	X	X	X	X	X	
	ETN	Drug	X	X	X			X
		Antibody	X	X	X			X
	IFX	Drug	X	X	X	X	X	X
		Antibody	X	X	X	X	X	
	GLM	Drug	X	X	X			X
		Antibody	X	X	X			
	CTZ	Drug			X			X
		Antibody			X			
Secondary non-responder	ADL	Drug	X	X	X	X	X	X
		Antibody	X	X	X	X	X	
	ETN	Drug	X	X	X			X
		Antibody	X	X	X			X

Clinical scenario	TNF-α inhibitor	Drug/Antibody	ELISAs					
			Promonitor	IDKmonitor	LISA-TRACKER	RIDASCREEN	MabTrack	Sanquin
	IFX	Drug	X	X	X	X	X	X
		Antibody	X	X	X	X	X	
	GLM	Drug	X	X	X			X
		Antibody	X	X	X			
	CTZ	Drug			X			X
		Antibody			X			

Key: ADL: adalimumab; CTZ: certolizumab pegol; ETN: etnercept; IfX: infliximab; GLM: golimumab

As previously noted, and as seen in Table 13, the technologies will be evaluated i) for use during remission/LDA to inform decisions regarding whether the same treatment should continue at the same dose; ii) to identify primary non-responders; iii) to identify and examine potential reasons for secondary non-response.

2 Assessment of clinical effectiveness

This review assessed the clinical effectiveness of using enzyme-linked immunosorbent assay (ELISA) tests for measuring drug levels (adalimumab [ADL], etanercept [ETN], infliximab [IFX], certolizumab pegol [CTZ] and golimumab [GLM]) and/or their anti-drug antibodies (anti-ETN, anti-IFX, anti-ADL, anti-CTZ and anti-GLM) for the purpose of monitoring response to those tumour necrosis factor-alpha (TNF-α) inhibitors in people with rheumatoid arthritis (RA). The eligible populations were people with RA who were being treated with TNF-α inhibitor therapies and:

- had achieved treatment target (remission or low disease activity [LDA]) or,
- experienced a primary non-response or,
- experienced a secondary non-response.

2.1 Methods for reviewing effectiveness

The systematic review was conducted following the Cochrane Handbook for Systematic Reviews of Intervention[54] and the PRISMA statement.[55] We performed the systematic review according to a pre-specified protocol which was registered on the international prospective register of systematic reviews (PROSPERO: CRD 42018105195).

2.2 Identification of studies

The following bibliographic databases were searched :

- MEDLINE (Ovid)
- MEDLINE In-Process & Other Non-Indexed Citations (Ovid)
- Embase (Ovid)
- Web of Science (Thomson Reuters)
- Cochrane Database of Systematic Reviews, CENTRAL (via the Cochrane library).

In addition, searches were carried out on the following websites: Proquest theses, British Library theses, DART-Europe theses, Prospero, ARIF, HTA database, DARE, CRD, Open Grey, Grey literature report, C-EBLM, British Society for Rheumatology, EULAR, American Society for Rheumatology, Medion Grifols, Theradiag, Sanquin, R-Biopharm, Immunodiagnostic, Biohit, Progenika, Clinical Trials.gov, WHO Registry and EU trials register. The following resources: Clinical Trials.gov, WHO Registry and EU trials register provide coverage for ongoing trials.

The search strategies were developed by an information specialist in July 2018, and were designed to be as sensitive as possible. They comprised terms for RA and terms for anti-TNF inhibitors and terms for ELISA testing. No study type, language or date filters were used; studies were limited to human only (not animal studies) where appropriate. The search was conducted in late July 2018. An updated search was performed on 19 November 2018.

The full search strategies for each database are reported in Appendix 1. The search results were exported to Endnote X7 (Thomson Reuters, NY, USA) and deduplicated using automatic and manual checking.

Items included after full-text screening were forward and backward citation chased using Scopus (Elsevier) in order to identify additional relevant studies. The reference lists of potentially relevant systematic reviews were checked for additional relevant studies. The references lists that were submitted by industry were also checked in order to identify additional relevant studies.

2.2.1 Inclusion and exclusion criteria

The inclusion criteria for the clinical effectiveness review were as follows:

2.2.1.1 Population

The eligible population was people with RA receiving treatment with a TNF-α inhibitor (ADL, ETN, IFX, CTZ and GLM), and:

- had achieved treatment target (remission or LDA) or,
- experienced a primary non-response or,
- experienced a secondary non-response.

2.2.1.2 Interventions

ELISA test kits or diagnostic services used to monitor response to TNF-α inhibitor treatments for people with RA were eligible for inclusion. These tests run on an ELISA technology platform, and are used to measure drug levels (ADL, ETN, IFX, CTZ, and GLM) or their anti-drug antibodies (anti-ETN, anti-IFX, anti-ADL, anti-CTZ, and anti-GLM). A serum sample is needed to perform an ELISA test.

Eligible ELISA tests can be run with or without automation platforms and may be used with any ELISA platform or the Tritutus and SQII platforms. Each test only needs to be run once, potentially allowing for high throughput. The test should be intended for monitoring purpose to inform treatment decisions to biologic therapies in people with RA.

The ELISA testing kits or diagnostic services shown below were included:

Promonitor ELISA kits (Grifols-Progenika):

- Promonitor-ADL-1DV
- Promonitor-ANTI-ADL-1DV
- Promonitor-ETN-1DV
- Promonitor-ANTI-ETN-1DV
- Promonitor-GLM-1DV
- Promonitor-ANTI-GLM
- Promonitor- IFX-1DV
- Promonitor-ANTI-IFX-1DV

IDKmonitor ELISA kits (Immundiagnostik/BioHit Healthcare):

- IDKmonitor adalimumab drug level
- IDKmonitor adalimumab free ADA
- IDKmonitor adalimumab total ADA
- IDKmonitor etanercept drug level
- IDKmonitor etanercept free ADA
- IDKmonitor golimumab
- IDKmonitor golimumab free ADA
- IDKmonitor infliximab drug level
- IDKmonitor infliximab free ADA
- IDKmonitor infliximab total ADA

LISA-TRACKER ELISA kits (Theradiag):

- LISA-TRACKER Adalimumab (LTA002)
- LISA-TRACKER anti-Adalimumab (LTA003)
- LISA-TRACKER Duo Adalimumab (LTA005)
- LISA-TRACKER Certolizumab (LTC002)
- LISA-TRACKER anti-Certolizumab (LTC003)
- LISA-TRACKER Duo Certolizumab (LTC005)
- LISA-TRACKER Etanercept (LTE002)
- LISA-TRACKER anti-Etanercept (LTE003)
- LISA-TRACKER Duo Etanercept (LTE005)
- LISA-TRACKER Golimumab (LTG002)

- LISA-TRACKER anti-Golimumab (LTG003)
- LISA-TRACKER Duo Golimumab (LTG005)
- LISA-TRACKER Infliximab (LTI002)
- LISA-TRACKER anti-Infliximab (LTI003)
- LISA-TRACKER Duo Infliximab (LTI005)

RIDASCREEN ELISA kits (R-Biopharm)

- RIDASCREEN ADM monitoring
- RIDASCREEN anti-ADM antibodies
- RIDASCREEN IFX monitoring
- RIDASCREEN anti-IFX antibodies

MabTrack ELISA kits (Sanquin)

- MabTrack level adalimumab M2910
- MabTrack ADA adalimumab M2950
- MabTrack level infliximab M2920
- MabTrack ADA infliximab M2960

Sanquin Diagnostic Services (testing service using validated ELISAs)

- ADL drug levels
- CTZ drug levels
- ETN drug levels
- ETN anti-drug antibodies
- GLM drug levels
- IFX drug levels

The use of both free and total anti-drug antibody assays for these tests were assessed, depending on the availability of assessment data relating to both assays. The intervention tests were used in addition to current clinical practice (clinical assessment and monitoring using a composite score such as disease activity score in 28 joints (DAS28).

2.2.1.3 Comparator

Standard care for people with RA where treatment decisions were based on clinical judgements and monitoring using a composite score such as the disease acticivty score 28 joints (DAS28), without the knowledge of circulating drug levels and anti-drug antibodies by ELISA tests.

2.2.1.4 Outcomes

There was no restriction on when the outcomes were measured. The following outcomes were included:

- Test (procedural) outcomes
    - Number of inconclusive test results
    - Time to test result
- Treatment and management outcomes:
    - Number, direction and magnitude of dose changes
    - Frequency of dose adjustment (e.g. dose reduction) due to monitoring response
    - Frequency of treatment switch to an alternative biologic
    - Discontinuation of ineffective therapy
- Clinical outcomes:
    - Change in disease activity
    - Rates of disease response, relapse and remission
    - Duration of response, relapse and remission
    - Rates of hospitalisation
    - Rates of surgical intervention
    - Adverse effects (AEs) of treatment such as infections
- Patient-related outcomes
    - Health related quality of life (HRQoL).

The primary clinical outcomes were clinical and patient-related outcomes including improvement on disease activity and HRQoL. The clinically important intermediate outcomes were change in number, direction and magnitude of anti-TNF dose, change in frequency of dose adjustment due to monitoring response, change in frequency of treatment switch to an alternative biologic, and discontinuation of ineffective therapy.

2.2.1.5 Study design

Both randomised controlled trials (RCTs) and non-randomised controlled studies comparing therapeutic drug monitoring (TDM) by using ELISA tests with standard care were included. Observational studies that evaluated the clinical effectiveness of the intervention tests to monitor treatment response in pe0ople with RA were included, providing they reported any of those relevant clinical outcomes for this assessment. Examples of observational studies included: prospective cohort studies, retrospective cohort studies and historically controlled studies.

2.2.1.6 Exclusions

The following types of report were excluded: editorials and opinions; case reports; reports focusing only on technical aspects of the technologies (such as technical descriptions of the testing process). Non-English studies were excluded. Studies with a sample size of 20 or less were excluded due to inadequate statistical power. For studies that included people with RA, ankylosing spondylitis (AS) and psoriatic arthritis (PsA), studies with at least 70% of people with RA were included providing other eligibility criteria were met. The relevance of any studies that included less than 70% people with RA was consulted on with clinical experts, and study authors were contacted to try and get subgroup data for people with RA. In cases where there were multiple reports for a given study or when the possibility of overlapping populations could not be excluded the most recent or most complete report was selected.

2.2.2 Study selection strategy

Two reviewers screened independently the titles and abstracts (if available) of all reports identified by the search strategy. Full text copies of all studies deemed to be potentially relevant were obtained and two reviewers independently assessed them for inclusion. Any disagreements were resolved by consensus.

2.2.3 Data extraction strategy

A data extraction form was developed and piloted. One reviewer independently extracted details of study design, participants, interventions and outcome data. The data extraction was checked by another reviewer. Any disagreements were resolved by consensus.

For studies reporting clinical event outcomes data were extracted on these as numbers of people experiencing the specified outcome. For studies reporting continuous outcomes we extracted data on these as mean and standard deviation. Where reported, mean differences, relative risks, odds ratios or incidence rate ratios (with 95% confidence intervals) were extracted from comparative studies. Where available, results adjusted for potential confounding factors (such as age, gender and disease duration of rheumatoid arthritis) were extracted preferentially.

For studies in which only a subgroup of people were eligible for inclusion in the review, data were extracted and presented for this subgroup only. If some data were unclear or missing, attempts were made to contact the study authors to obtain additional data.

2.2.4 Critical appraisal strategy

One reviewer independently assessed the quality of included studies in terms of risk of bias. If RCTs had been identified, the Cochrane Risk of Bias tool for RCTs would have been used.[56] The Cochrane (ROBINS-1) tool was used for non-randomised studies with adaptations as appropriate.[57] We also used the Cochrane (ROBINS-1) tool to assess the quality of uncontrolled observational studies with adaptations as appropriate, although the tool was primarily designed for non-randomised controlled studies. The risk of bias of included studies was taken into account when interpreting results. The quality assessment was checked by another reviewer. Any disagreements were resolved by consensus.

2.2.5 Methods of data synthesis

Given the clinical heterogeneity associated with interventions, outcomes and length of follow-up and the methodological heterogeneity identified (e.g. different study designs), quantitative synthesis was not possible and clinical effectiveness data were synthesised in a narrative fashion. Publication bias could not be investigated because quantitative synthesis was not possible.

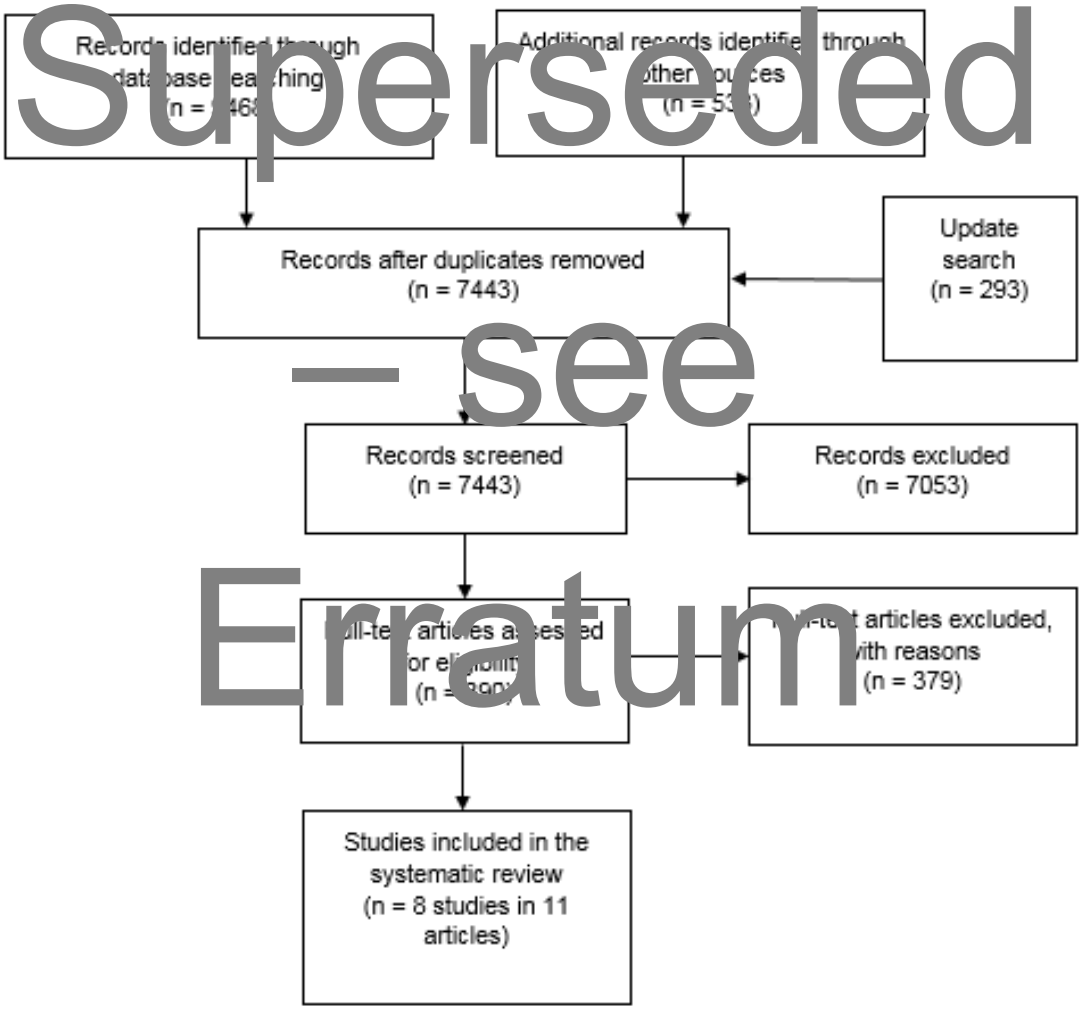
2.3 Clinical effectiveness results

The next section provides information on the quantity of research available, including characteristics and risk of bias of the included studies. This is then followed by the results section with clinical effectiveness of therapeutic drug monitoring by using ELISA tests in people with RA who were treated with TNF-α inhibitors.

2.3.1 Quantity and quality of research available

The literature searches of bibliographic databases identified 7,443 references. After initial screening of titles and abstracts, 390 were considered to be potentially relevant and were ordered for full paper screening. In total, eight studies reported in 11 articles[11,12,15,58-65] were included in the systematic review of clinical effectiveness of using ELISA tests for monitoring response to TNF-α inhibitors in people with RA. All the included trials with linked citations are presented in Appendix 2. Figure 2 shows a flow diagram outlining the screening process with reasons for exclusion of full-text papers.

Figure 2: Flow diagram of study inclusion process for the clinical-effectiveness review



Some studies were reported in multiple sources and abstracts, with considerable overlap in data and reporting. The paper with the most up-to-date and complete data was selected for the data extraction.

Number and type of studies excluded

A list of full-text papers that were excluded along with the reasons for their exclusions is given in Appendix 3. These papers were excluded because they failed to meet one or more of the inclusion criteria in terms of the type of study design, participants, interventions or outcomes being reported.

2.3.2 Assessment of clinical effectiveness

2.3.2.1 Characteristics of included studies

The characteristics of included studies are presented in Table 14 and Table 15.[11,12,15,58-65] Most studies recruited people with RA who had achieved treatment target (remission or

LDA). Only one study[62] recruited people with RA who had experienced a primary non-response or a secondary non-response.

Five studies used Promonitor ELISA kits for monitoring drug levels and/or anti-drug antibody levels.[11,12,15,58,63-65] Three studies in four sources[59-62] used Sanquin ELISA kits to measure drug levels and/or anti-drug antibody levels of three anti-TNFs (IFX, ADL, and ETN for the treatment of RA. The type of Sanquin test kits used in these studies was not reported. Seven studies were conducted in Spain while only one study[65] was conducted in Asia (Taiwan). Two studies (reported in three publications) (Inciarte-Mundo and colleagues, 2016; Paredes and colleagues 2015; Paredes and colleagues 2016) were sponsored by pharmaceutical companies but other studies did not state funding sources.[60,61,63]

Non-randomised controlled studies

Three abstracts[11,15,64] were identified reporting the same non-randomised control trial (the INGEBIO study). In this trial, monitoring testing results of drug levels and anti-drug antibodies were revealed to physicians in the intervention arm. The monitoring test results were not revealed to physicians in the control arm. This reflected standard care in Spain where treatment decisions were based on clinical judgements without the knowledge of drug levels and anti-drug antibodies. Given that this was a pragmatic trial, it is likely that the findings could be generalisable to routine practice settings. For standard care in the control arm, clinicians did not follow any national guideline for the management of people with RA as there were no national guidelines for monitoring in Spain at the time of the study. Clinicians used their best judgements to optimise treatment doses. This trial recruited a mixed population of 169 people with RA (n=63), PsA (n=54) and ankylosing spondylitis (n=52) recruited from three sites in Spain. The study focused on the population who had achieved treatment target (remission or LDA) and remained clinically stable for at least six months.

The included abstracts reported a sample size of people with RA ranging from 54 to 63 at baseline. The abstracts by Ucar and colleagues (2017) and Arango and colleagues (2017) reported results on the basis of 18-month follow-up. The abstract by Gorostiza and colleagues[64] reported results only based on 34-week follow-up. This trial reported the following relevant clinical outcomes: change in disease response, dose adjustment due to monitoring response (e.g. proportion of participants tapered), and participants' HRQoL outcomes.

The median duration of disease at baseline among participants in the three abstracts[11,15,64] ranged from 117 to 124 months. All participants were treated with adalimumab (ADL) 40 mg

(via subcutaneous injection). ADL and anti-adalimumab antibody (ADAb) levels were measured using Promonitor-ADL and Promonitor-ANTI-ADL (Grifols-Progenika). The frequency of testing in this trial was once every two to three months. There were a total of eight visits during the trial period (details were not provided).

Observational studies

Seven observational studies reported in eight articles[12,58-63,65] assessed the clinical effectiveness of using ELISA tests for monitoring response to TNF-α inhibitors in people with RA. Paredes (2015)[60] and Paredes (2016)[61] reported the same study at different follow-ups. Most observational studies recruited people who had achieved treatment target (remission or LDA). Only one study[62] recruited people who had experienced a primary non-response or a secondary non-response. One observational study[59] had a historical control while the remaining observational studies were single-arm trials with no comparator. These observational studies reported the following relevant clinical outcomes: change in disease response, change in disease activity, change in direction and magnitude of therapeutic dose, and discontinuation of ineffective therapy.

The majority of observational studies used Promonitor ELISA kits for monitoring drug levels and/or anti-drug antibody levels. Three observational studies in four articles (51-54) used Sanquin ELISA kits for measuring drug levels for three anti-TNFs (IFX, ETN, ADL). The sample size of included studies ranged from 36 to 64. Where reported, the frequency of measurement varied between included studies. The majority of observational studies measured drug levels and/or anti-drug antibody levels once every four to six months.

Only three observational studies[12,59,63] measured anti-TNF drug levels only. The majority of included studies measured both anti-TNF drug levels and anti-drug antibody levels. However, it was unclear whether the drug levels and anti-drug antibody levels were assessed concurrently as the studies did not report the testing method (such as concurrent testing and reflex testing). In studies where anti-drug antibody levels were measured, it was unclear whether free anti-drug antibody assays or total anti-drug antibody assays were assessed. For studies measuring drug levels, only two studies[63,65] reported that serum drug trough levels were measured by ELISA tests. It was unclear whether drug trough levels were assessed in the remaining studies.

The included studies did not report other outcomes such as number of inconclusive results, time to result, frequency of treatment switch to an alternative biologic, rates of hospitalisation and rates of surgical interventions.

These observational studies measured drug levels and/or anti-drug antibody levels in participants treated with ADL, ETN and/or IFX. No studies were identified in which participants were treated with certolizumab pegol and golimumab. No studies reporting on the use of ELISA testing in people with RA receiving biosimilar products were identified. No relevant studies (including both controlled trials and observational studies) were identified that assessed other eligible ELISA kits including IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits.

Table 14: Characteristics of the included studies – The INGEBO non-randomised controlled study

Study[a]	Location	Population	Sample size	Median disease duration (months	Description of tests	Description of intervention	Description of control	Length of follow-up (months)	Number of visits
Ucar 2017[11]	Spain	Remission/LDA for at least 6 months	169[b]	117	ADL/ADAb serum levels using Promonitor-ADL and Promonitor-ANTI-ADL (Grifols - Progenika).	40mg subcutaneous ADL; TDM data released to physician	40mg subcutaneous ADL; TDM data not released to physician	18	8
Gorostiza 2016[64]	Spain	Remission/LDA for at least 6 months	169[c]	117	ADL/ADAb serum levels using Promonitor-ADL and Promonitor-ANTI-ADL (Progenika, Spain)	40mg subcutaneous ADL; TDM data released to physician	40mg subcutaneous ADL; physician blinded to TDM data	34-week	8
Arango 2017[15]	Spain	Remission/LDA	169[d]	124	ADL/ADAb serum levels using Promonitor-ADL and Promonitor-ANTI-ADL (Grifols - Progenika).	40mg subcutaneous ADL; TDM data released to physician	40mg subcutaneous ADL; TDM data not released to physician	18	8

Key: ADL: adalimumab; ADAb: anti-adalimumab antibody; LDA: low disease activity; NR: not reported, TDM: therapeutic drug monitoring

Notes:

[a] Study date not reported

[b] Sample size for people with RA was 63

[c] Sample size for people with RA was 63

[d] Sample size for people with RA was 54

Table 15: Characteristics of the included studies – Observational studies

Study	Study date	Location	Study design	Population	Description of tests	Frequency of Measuring	Sample size	Length of follow-up
Pascual-Salcedo 2013[59]	2006-2012	Spain	Historically controlled study	Remission	Drugs: IFX, ADL, ETN Capture ELISA (Sanquin, Amsterdam)	NR	43	7 years
Senabre 2017[12]	2011-2016	Spain	Prospective uncontrolled cohort study	Clinical remission	Serum ADL and ETN drug levels measured with Promonitor ELISA kits (Progenika)	6 monthly	39	1 year
Chen 2016[65]	NR	Taiwan	Prospective uncontrolled cohort study	Clinical remission/LDA	Anti-ADAb measured using bridging ELISA; serum ADL trough levels measured using sandwich ELISA (Progenika Biopharma, Spain)	At baseline and at 24 weeks	64	24 weeks
Inciarte-Mundo 2016[63]	NR	Spain	Prospective uncontrolled cohort study	Remission/LDA for ≥3 months	Serum trough levels of ADL, ETN or IFX measured with Promonitor ELISA kits (Progenika)	Every 4 months; and at disease flare	47	1 year
Lopez-Casla 2013[62]	2000	Spain	Prospective uncontrolled cohort study	Primary and secondary non-responders	Serum drug (IFX) measured using capture ELISA; anti-IFX antibody measured using bridging ELISA (Sanquin, Amsterdam)	Baseline, before increasing dose, at 6 months and 1 year	23/36 primary non-responders; 13/36 secondary non-responders	1 year
Rosas 2015[58]	2013-2014	Spain	Prospective uncontrolled cohort study	Remission for 6 consecutive months	Serum drug and anti-drug antibodies levels were measured using promonitor-ADL, promonitor-ETN, promonitor-Anti-ADL and promonitor-Anti-ETN (Progenika, Grifols, Spain)	Before each injection	45	NR
Paredes 2015[60]	NR	Spain	Retrospective uncontrolled cohort study	Remission/LDA for at least 6 months	Serum drug and anti-drug antibodies were measured for ADL, IFX, ETN Capture ELISA (Sanquin, Amsterdam)	Measured at baseline and last available visit during the 2-year follow-up	54	2 years

Study	Study date	Location	Study design	Population	Description of tests	Frequency of Measuring	Sample size	Length of follow-up
Paredes 2016[61]	NR	Spain	Retrospective uncontrolled cohort study	Remission/LDA for at least 6 months	Serum drug and anti-drug antibodies were measured for ADL, IFX, ETN Capture ELISA (Sanquin, Amsterdam)	Measured at baseline before tapering (pre-visit) and last available visit after 4 years follow-up (final visit)	52	4 years

Key: ADL: adalimumab; ELISA: enzyme linked immunosorbent assay; ETN: etanercept; IFX: infliximab; LDA, low disease activity; NR, not reported

2.3.2.2 Baseline characteristics of included studies

Baseline characteristics of included studies are presented in Table 16 and Table 17. The mean age of participants enrolled across studies ranged from 53 to 61 years. The proportion of females was at least 75% of the total population in each study. Where reported, the mean disease duration of RA ranged from nine to 17 years across studies.

Where reported, the definitions of remission, low LDA and flare used were generally consistent between included studies (see Table 16 and Table 17). All studies used one or more anti-TNF therapies (ADL, IFX, or ETN) for the treatment of RA. The mean treatment duration for participants receiving anti-TNF inhibitors ranged from three to six years.

Where reported, the included studies used different types of co-therapies for the management of people with RA. Co-therapies included methotrexate, sulfasalazine, hydrochloroquine, steroids (e.g. prednisolone), leflunomide, corticosteroids and other DMARDs).

Table 16: Baseline characteristics of included studies – The INGEBO* non-randomised controlled studies

Study/Author	Mean age	Definition of remission	Definition of LDA	% male	% remission at baseline	% LDA at baseline	N	Median disease duration (months)	Mean time on biologic (years)	Co-therapies	Anti-TNF received	Dose manipulation
Ucar 2017[11]	53.59[a]	NR	NR	58	70.0 (IG: 73.4, CG: 83.3)[b]	30.0 (IG: 26.6, CG: 16.7)[b]	169	117.0	NR	Methotrexate	ADL 40mg subcutaneous	Dose tapering; physicians alter dose based on their judgement
Gorostiza 2016[64]	NR	NR	NR	NR	70.0 (IG: 73.4, CG: 83.3)[b]	30.0 (IG: 26.6, CG: 16.7)[b]	169	117.0	NR	NR	ADL 40mg subcutaneous	Dose tapering; physicians alter dose based on their judgement
Arango 2017[15]	NR	NR	NR	NR	67.3 (IG: 71.4, CG: 82.7)[b]	32.7 (IG: 28.6, CG: 17.3)[b]	169	124.0	NR	NR	ADL 40mg subcutaneous	Dose tapering; physicians alter dose based on their judgement

Key: ADL: Adalimumab; CG: control group; NR: not reported; IG: intervention group; RA: rheumatoid arthritis

Notes:

[*] Same study reported in three abstracts

[a] Weighted mean across arms

[b] Percentages are reported for the combined population of RA, psoriatic arthritis and ankylosing spondylitis

Table 17: Baseline characteristics – observational studies

Study/ Author	Mean age (yrs)	Sample size	Definition of Remission (DAS28)	Definition of LDA (DAS28)	Definition of flare (DAS28)	% male	Mean disease duration (yrs)	Mean time on biologic (yrs)	Co-therapies	Anti-TNF received	Dose manipulation
Pascual-Salcedo 2013[59]	NR	43	< 3.2[c]	< 3.2[c]	NR	NR	17.52 (SD 9.38)	5.85 (SD 1.33)	NR	ADL; ETN; IFX (doses NR)	Optimization strategy (adjusting drug dose according to clinical activity)
Senabre Gallego 2017[12]	61 (range 39-81)	39	NR	NA	NA	18	14.95 (range 2.15-52.31)	4.21 (range 1.39-11.07)	Methotrexate, Leflunomide, Hydrochloroquine, Sulfasalazine, low-dose corticosteroid (doses NR)	ADL; ETN (dose NR)	Dose reduction by extension of anti-TNFα administration interval
Chen 2016[65]	55.45[a]	64	< 2.6	< 3.2	≥ 3.2	9.4	9.11[a]	2.89[a]	Methotrexate, Sulfasalazine, Hydrochloroquine (doses NR)	ADL 40mg monthly (route NR)	Tapering, dose-halving and monitoring for 24 weeks
Inciarte-Mundo 2016[63]	57 (range 30-81)	47	< 2.6	< 3.2	> 3.2	NR	NR	5.08	Steroids (type and doses NR)	ADL; ETN; IFX (doses NR)	Tapering (47% on reduced dose)
Lopez-Casla 2013[62]	58 (SD 3.6)	36	NA (non-responders)	NR	NR	NR	19.2 (SD 10.5)	6.6 (SD 3.8)	Methotrexate, other DMARDs and Prednisolone (doses NR)	IFX 3-5 mg/kg intravenous	Dose increase from minimum to maximum according to response
Rosas 2015[58]	60.5 (SD 18)	45	≤ 2.6	NA	NA	13	15 (SD 9.8)	ADL 5.1 (SD 1.3) ETN 5.1 (SD 1.8)	Synthetic DMARDs (type and doses NR)	ADL; ETN (dose NR)	Dose reduction by decreasing treatment frequency

Study/ Author	Mean age (yrs)	Sample size	Definition of Remission (DAS28)	Definition of LDA (DAS28)	Definition of flare (DAS28)	% male	Mean disease duration (yrs)	Mean time on biologic (yrs)	Co-therapies	Anti-TNF received	Dose manipulation
Paredes 2015[b60]	60.2 (SD 12)	54	<2.6	< 3.2	NR	22.2	NR	NR	Methotrexate, other DMARDs and prednisolone (doses NR)	ADL; ETN; IFX;(doses NR)	Optimisation strategy (tapering or increase in interval of administration according to response)
Paredes 2016[b61]	NR	52	<2.6	< 3.2	>3.2	21	NR	NR	Methotrexate, other DMARDs and prednisolone (doses NR)	ADL; ETN; IFX; (doses NR)	Tapering involving dose reduction or discontinuation

Key: ADL: adalimumab; DMARDs: disease-modifying anti-rheumatic drugs; ETN: etanercept; IFX: infliximab; LDA: low disease activity; NA: not applicable; NR: not reported; SD: standard deviation; TNF: tumour necrosis factor; yrs: years

Notes:

[a] Weighted

[b] Same study, different follow-up time

[c] Grouped as 'remission or LDA

2.3.2.3 Ongoing studies

One ongoing RCT was identified that met inclusion criteria for this systematic review of clinical effectiveness: the Norwegian Drug Monitoring Study (NOR-DRUM).[66] Study characteristics are summarised in Table 18. Enrolment for the NOR-DRUM trial commenced in March 2017, with expected primary completion date of March 2020 and study completion date of March 2022.

The aim of this trial is to assess the clinical effectiveness of TDM in participants starting IFX and in participants on maintenance IFX therapy. The type of ELISA testing is not reported. The target recruitment for this study 600 people with RA, and those with other immunological inflammatory diseases.

The intervention of this trial will be TDM with a treatment algorithm based on measurement of serum drug levels and anti-drug antibodies. The control group is standard care where clinicians will make treatment decisions without the knowledge of drug levels or status of anti-drug antibodies.

The major primary outcomes are the proportions of participants in remission and the proportions of participants in sustained disease control without disease worsening. Secondary outcomes of interest include time to sustained remission, occurrence of drug discontinuation, health utility (EQ-5D), HRQoL (SF-36), time to disease worsening and clinical efficacy outcomes assessed by composite disease activity scores.

Table 18: Characteristics of the Norwegian Drug Monitoring Study (NOR-DRUM)[66]

Study title	-	**A Norwegian Multicentre Randomised Controlled Trial Assessing the Effectiveness of Tailoring IFX Treatment by Therapeutic DRUg Monitoring - The NOR-DRUM Study**
Study objectives	-	Effectiveness of TDM in participants starting IFX
	-	Effectiveness of TDM in participants on maintenance IFX
Immunological inflammatory diseases enrolled	-	RA
	-	Spondyloarthritis
	-	Ankylosing spondylitis
	-	Crohn's disease
	-	Ulcerative colitis
	-	Psoriasis
	-	PsA
Intervention arm	-	TDM*
Comparator arm	-	Standard care**
N (expected)	-	600
Start date	-	March 1, 2017
Estimated primary completion date	-	March 1, 2020
Estimated study completion date	-	March 1, 2022
Outcomes	Primary	

Study title	-	**A Norwegian Multicentre Randomised Controlled Trial Assessing the Effectiveness of Tailoring IFX Treatment by Therapeutic DRUg Monitoring - The NOR-DRUM Study**
	-	Proportion of participants in remission defined by disease specific composite scores
	-	Sustained disease control throughout the study period without disease worsening defined by disease specific composite scores
	Secondary	
	-	Time to sustained remission
	-	Patient's and physician's global assessment of disease activity
	-	Change in ESR
	-	Change in CRP
	-	Occurrence of anti-drug antibodies
	-	Reason for drug discontinuation
	-	Occurrence of drug discontinuation
	-	Cost effectiveness
	-	Health utility (EQ-5D)
	-	Quality of life (SF-36)
	-	Safety (adverse events frequency)
	-	Efficacy assessed by composite disease activity scores
	-	Time to disease worsening
Eligibility criteria	NOR-DRUM A	
	-	A clinical diagnosis of one of the following; RA, spondyloarthritis (including ankylosing spondylitis), PsA*, ulcerative colitis, Crohn's disease or chronic plaque psoriasis
	-	Male or non-pregnant female
	-	≥18 and < 75 years of age at screening
	-	A clinical indication to start IFX
	-	Subject not in remission according to diagnosis-specific disease activity scores
	-	Subject capable of understanding and signing an informed consent form
	NOR-DRUM B	
	-	A clinical diagnosis of one of the following; RA, spondyloarthritis (including ankylosing spondylitis), PsA*, ulcerative colitis, Crohn's disease or chronic plaque psoriasis
	-	Male or non-pregnant female
	-	≥18 and < 75 years of age at screening
	-	On maintenance therapy with IFX for a minimum of 30 weeks and a maximum of 3 years
	-	A clinical indication for further IFX treatment

Key: IFX: inflixmab; PsA: psoriatic arthritis; RA: rheumatoid arthritis; TDM: therapeutic drug monitoring

Notes:

* Administration of IFX according to a treatment strategy based on TDM and assessments of anti-drug antibodies; treatment algorithm based on assessments of serum drug levels and anti-drug antibodies

** Administration of IFX according to standard clinical care, without knowledge of drug levels or status of anti-drug antibodies; treatment algorithm based on standard clinical assessments, without knowledge of serum drug levels and anti-drug antibodies

2.3.2.4 Risk of bias of included studies

The risk of bias of included studies was assessed using the Cochrane (ROBINS-1) tool for non-randomised studies. The Cochrane (ROBINS-1) tool was also used to assess the quality of uncontrolled observational studies with adaptations as appropriate, although the tool was primarily designed for non-randomised controlled studies. The following domains

relating to risk of bias were assessed for each individual study: confounding, selection, group classification, co-interventions, missing data, outcome measurement and selective outcome reporting. The quality assessments on the basis of all relevant domains for each study and of specific outcomes are presented in Appendix 3, Table 17 and Table 20 present the quality assessment of included studies.

Table 19 presents the quality assessment of the non-randomised controlled study (the INGEBIO study).[11,15,64] This non-randomised controlled study was judged to be at moderate risk of bias. There was an issue of baseline imbalance in the proportions of participants with remission and LDA between the intervention and control groups: 73.4% of participants were in remission at baseline in the intervention group while 83.3% of participants were in remission at baseline in the control group. The remaining participants (i.e., 26.6 % of participants in the intervention group and 16.7% of participants in the control group) had achieved LDA at baseline. Furthermore, there was a lack of adjustment for this baseline imbalance variable in the analysis of clinical outcomes. These deficiencies resulted in serious risk of bias associated with the findings.

Table 21 presents the attrition rates for each outcome of the non-randomised controlled study (the INGEBIO study).[11,15,64] As seen in Table 21 there were high attrition rates for three outcomes (proportions of participants who remained in remission, proportions of participants who changed from LDA to remission and proportions of participants who received dose tapering). The attrition rates ranged from 11.2% to 30.8%, which can lead to attrition bias. Furthermore, there were unbalanced attrition rates in these outcomes between the intervention and control groups.

Table 20 presents the quality assessment of observational studies. Among all observational studies, only one study[59] had a historical control group but other studies were single arm studies with no comparator group. The study by Pascual-Salcedo (2013)[59] was judged to be at moderate risk of bias because there was non-contemporaneous control bias due to the use of historical control in this study. It should be noted that the same group of participants were assessed during the first period (i.e. the historical control where TDM was not introduced) and the second period (where TDM was implemented).

All single arm studies[12,58,60-63,65] were judged to be at moderate risk of bias. Across these studies, there were low to moderate risks of biases in the domains of confounding, selection, group classification, co-interventions, missing data, outcome measurement and selective outcome reporting. Therefore, these studies were deemed to be at moderate risk of bias. In

particular, two studies[61,62] had an issue of missing data, with the attrition rates ranging from 3.7% to 5.5% (see Table 22).

Overall, the non-randomised controlled study[11,15,64] was judged to be at serious risk of bias. For observational studies, the historically controlled study and all the single arm studies were judged to be at moderate risk of bias.

Table 19: Risk of bias in included studies – The INGEBIO non-randomised controlled study

Studies	Confounding (differential prognosis between groups)	Selection	Group Classification	Co-intervention	Missing data	Outcome measurement	Selective outcome reporting	Overall risk of bias
Arango 2017[15]	Serious	Low	Low	NI	Serious	Moderate	Low	Serious
Gorostiza 2016[64]	Serious	Low	Low	NI	Serious	Moderate	Low	Serious
Ucar 2017[11]	Serious	Low	Low	NI	NI	Moderate	Low	Serious

Risk of bias judgement: low/moderate/serious/critical/NI NI: no information

Table 20: Risk of bias in included studies – observational studies

Studies	Confounding	Selection	Group Classification	Co-intervention	Missing data	Outcome measurement	Selective outcome reporting	Overall risk of bias
Chen 2016[65]	NA*	Low	Low	Moderate	Low	Moderate	Low	Moderate
Inciarte-Mundo 2016[63]	NA*	Low	Moderate	Moderate	NI	Moderate	Low	Moderate
Lopez-Casla 2013[62]	NA*	Moderate	Low	Moderate	Moderate	Moderate	Low	Moderate
Paredes 2015[60]	NA*	Low	Low	NI	NI	Moderate	Low	Moderate
Paredes 2016[61]	NA*	Low	Low	NI	Moderate	Moderate	Low	Moderate
Pascual-Salcedo 2013[59]	Moderate	Moderate	Moderate	NI	NI	Moderate	Low	Moderate
Rosas 2015[58]	NA*	Moderate	Low	NI	NI	Moderate	Low	Moderate
Senabre Gallego 2017[12]	NA*	Low	Low	Moderate	NI	Low	Low	Moderate

Key: Risk of bias judgement: low/moderate/serious/critical/NI NI: no information, NA*: not applicable because of lack of comparator group (tool not originally designed for single-arm studies)

Table 21: Attrition in the INGEBIO non-randomised controlled study

Outcome	Baseline population		Follow-up population		Percent attrition		
	IG	CG	IG	CG	IG	CG	Overall
Disease flare	109	60	Unclear	Unclear	Indeterminate	Indeterminate	Indeterminate
% remaining in remission	109	60	71	46	34.9 (109-71/109)	23.3 (60-46/60)	30.8 (169-117/169)
% change from LDA to remission	29	10	28	7	3.4 (29-28/29)	30.0 (10-7/10)	11.4 (39-35/35)
ADL tapering	109	60	98	52	10.1 (109-98/109)	13.3 (60-52/60)	11.2 (169-150/169)
HRQoL	109	60	Unclear	Unclear	Indeterminate	Indeterminate	Indeterminate

Key: ADL: adalimumab; CG: control group; HRQoL: health-related quality of life; IG: intervention group; LDA: low diease activity

Table 22: Attrition in observational studies

Outcomes	Study	Baseline population	Follow-up population	Percent attrition
Worsening of clinical activity	Senabre Gallego 2017[12]	39	NI	Indeterminate
Persistent remission	Chen 2016[65]	64	64	0
Persistent LDA	Chen 2016[65]	64	64	0
Remission turned LDA	Chen 2016[65]	64	64	0
Disease flare	Chen 2016[65]	64	64	0
Disease flare	Inciarte-Mundo 2016[63]	47	NI	Indeterminate
Modified dosing frequency	Rosas 2015[58]	45	NI	Indeterminate
Total doses avoided	Rosas 2015[58]	45	NI	Indeterminate
Disease flare	Paredes 2015[60]	54	NI	Indeterminate
Pre-visit/final visit remission	Paredes 2016[61]	52	52	3.7% (2/54)*
Mean drug levels	Paredes 2016*[61]	52	NI	Indeterminate
Mean DAS scores	Pascual-Salcedo 2013[59]	43	NI	Indeterminate
Weekly mean drug dose	Pascual-Salcedo 2013[59]	43	NI	Indeterminate
Mean interval of drug administration	Pascual-Salcedo 2013[59]	43	NI	Indeterminate
Treatment discontinuation	Lopez-Casla 2013[62]	36	34	5.5% (2/36)**
AE (septic arthritis)	Senabre Gallego 2017[12]	39	NI	Indeterminate

Key: AE: adverse event; DAS: disease activity score; LDA: low disease activity; NI: no information

Notes:

* Paredes and colleagues (2016)[61] incldued a population of 52 participants appeared to be a four-year follow up of the two-year Paredes and colleagues (2015) study (54 participants were enrolled in the Paredes and colleagues (2015)[60] study)

** 26 of 36 (baseline denominator) = 72.2%, but this was reported as 76.5% implying two participants were not accounted for in the final analysis

2.3.3 Results of clinical effectiveness

2.3.3.1 Non-randomised controlled trials

Three included abstracts[11,15,64] reported the same non-randomised controlled trial (the INGEBIO study). This trial recruited participantswho had achieved treatment target (remission or LDA) and remained clinically stable for at least six months.

This trial recruited a mixed population of 169 participants including 63 people with RA. The results of the total mixed population were reported in the review, as the authors were not able to provide the results for the cohort of 63 people with RA (the study was not powered to detect a meaningful difference between the intervention and control groups for the cohort of people with RA only). The three cohorts of participants with different conditions (RA, PsA and AS) may have different treatment responses to TNF-α inhibitor therapies. Therefore, there was limited generalisability of findings from this mixed population to the RA population. At baseline, median trough levels of ADL were 5.3 mg/L for the intervention group and 5.5 mg/L for the control group. The quality of included abstracts was judged to be at serious risk of bias (see Section 2.3.2.4). Table 23 and Table 24 present the results of this non-randomised controlled study.

Change in disease response

The abstract by Ucar and colleagues (2017)[11] reported that at 18-month follow-up, the number of participants who had experienced a disease flare in the intervention and control groups was 69 and 53, respectively. In this study, a disease flare was defined as an increase in DAS28 >1.2 or >0.6 if DAS28 ≥3.2 following the criteria validated in the study by van der Maas and colleagues (2013).[67] As seen in Table 23, the rate of flares per patient-year is 0.463 for the intervention group and 0.639 for the control group, with rate difference of -0.176 (95% confidence interval (CI) -0.379 to 0.0289). There was a non-significant reduction in risk of flare in the intervention group compared with the control group (incidence rate ratio (IRR) 0.7252, 95% CI 0.4997 to 1.0578). Median time to first flare was 145 days in the intervention group and 136.5 days in the control group.

The number of participants who remained in remission at 18-month follow-up was not reported by Ucar and colleagues (2017);[11] however, the abstract by Gorostiza and colleagues 2016[64] reported that at 34-week follow-up, 76.1% (54/71) in the intervention group remained in remission while 69.6% (32/46) in the control group remained in remission. This analysis did not use an intention-to-treat (ITT) approach. By using the ITT analysis, the finding showed that 67.5% (54/80) in the intervention group remained in remission while

64.0% (32/50) in the control group remained in remission, with the difference in proportions of 3.5% (95% CI -13.3% to 20.3%; p=0.68).

This abstract (Gorostiza and colleagues 2016)[64] further reported that in participants with LDA at baseline, 35.7% (10/28) and 28.6% (2/7) were in remission at 34-week follow-up for the intervention and control groups, respectively. Again, this analysis did not use an ITT approach. Using the ITT analysis, the finding showed that in those participants with LDA at baseline, 34.5% (10/29) and 20% (2/10) were in remission at 34-week follow-up for the intervention and control groups, respectively.

Dose adjustment due to monitoring response

The abstract by Arango and colleagues 2017[15] reported that ADL dose was tapered in 35 participants in the intervention group (35.7%) and in 18 participants in the control group (34.6%). The results appeared to be generally similar between the intervention and control groups.

Health-related quality of life

Table 24 presents the results of HRQoL outcomes. Both Ucar and colleagues 2017[11] and Arango and colleagues 2017[15] reported the outcomes of participants' HRQoL (EQ-5D-5L). The results showed that participants' HRQoL outcomes (EQ-5D-5L) measures were higher in the intervention group at all visits compared with the control group (further details were not reported). However, statistically significant results were only observed at Visit 2 (p=0.001) and Visit 3 (p=0.035).

In summary, the findings from this non-randomised controlled trial (the INGEBIO study) showed that there was a non-significant reduction in risk of flare in the intervention group compared with the control group. Participants' HRQoL measures were higher in the intervention group at all visits compared with the control group, with statistically significant results being observed at two visits. However, given that the quality of this trial was judged to be at serious risk of bias, it may have compromised the reliability of the findings.

Table 23: Changes in disease response, relapse and remission

Study	Population	Intervention group (N)	Control group (N)	Length of follow-up (months).	Outcome measure	Relative measure[a]
Ucar 2017[11]	Remission/LDA	109	60	18	Number of disease flare	IG = 69, CG=53
					Incidence rate ratio (disease flare)	0.7252 (95% CI = 0.4997 to 1.0578)[e]
					Rate difference (disease flare)	-0.176 (95% CI = -0.379 to 0.0289)[e]
Gorostiza 2016[64]	Remission/LDA	109	60	18[b] (reported 34-week follow-up data)	% remained in remission[c]	69.6% (32/46) (CG) and 76.1% (54/71) (IG)
					Change from LDA to remission[d]	35.7% (10/28) (IG) and 28.6% (2/7) (CG)
Arango 2017[15]	Remission/LDA	98	52	18	% tapered	18/52 (34.6%) (CG), 35/98 (35.7%) (IG)
					Rate of flare	0.639 (CG), 0.463 (IG) flares/patient-year
					Rate difference	-0.176 (95% CI: -0.379 to 0.0289)[e]
					IRR	0.7252 (95% CI: 0.4997 to 1.0578)[e]
					Median time to 1st flare	136.5 (CG), 145 (IG) days

Key: CG: control group; IG: intervention group; IRR incidence rate ratio; LDA: low disease activity

Notes:

[a] The study population was mixed and included a total of 169 participants with RA, PsA, and ankylosing spondylitis

[b] 34-weeks follow-up results, as reported by authors

[c] Intention-to-treat (ITT) analysis; 67.5% (54/80) in IG and 64.0% (32/50) in CG remained in remission

[d] ITT analysis; 34.5% (10/29) (IG) and 20% (2/10) (CG)

[e] No specific number of patients for results specified

Table 24: Health-related quality of life outcomes

Study	Population	IG (N)	CG (N)	Length of follow-up (months)	Outcome measure	Relative measure[a]	P-value (IG vs CG) Visit 2	P-value (IG vs CG) Visit 3
Ucar 2017[11]	Remission/ LDA	109	60	18	Health-related quality of life (EQ-5D-5L)	Higher in IG throughout follow-up[b]	0.001	0.035
Arango 2017[15]	Remission/ LDA	98	52	18	Health-related quality of life (EQ-5D-5L)	Higher in IG throughout follow-up[b]	0.001	0.035

Key: CG: control group; IG: intervention group; LDA: low disease activity

Notes:

[a] All data included a mixed population of 169 patients (rheumatoid arthritis (RA), psoriatic arthritis (PsA) and ankylosing spondylitis)

[b] No specific number of patients for results specified

2.3.3.2 Observational studies

Eight observational studies[12,58-63,65] evaluated the effect of using ELISA tests for monitoring response to TNF-α inhibitors in people with RA.

Most observational studies included participants who had achieved treatment target (remission or LDA). Only one observational study[62] included participants who had experienced a primary non-response or a secondary non-response. Only one observational study[59] had a historical control while the remaining observational studies were single-arm studies with no comparator. The study by Lopez-Casla and colleagues (2013)[62] did not report the definitions of primary non-response and secondary non-response.

All the observational studies were judged to be at moderate risk of bias (see Section 2.3.2.4).

2.3.3.2.1 Change in disease response

Six observational studies assessed changes in disease response in people with RA who had achieved remission or LDA. The majority of studies were prospective uncontrolled cohort studies but two studies[60,61] were retrospective uncontrolled cohort studies. The sample size of included studies ranged from 39 to 64. Where reported, two studies[12,61] reported missing data at follow-up. Two studies[63,65] had no missing data in the follow-up period.

The duration of follow-up of the included studies ranged from 24 weeks to four years. All studies that assessed change in disease response focused on participants who had achieved treatment target (remission or LDA). Table 25 presents the results of changes in disease responses.

Four studies[60,61,63,65] evaluated the effect of optimisation of anti-TNF therapies (by decreasing dose or treatment frequency) guided by TDM in participants who had achieved remission or LDA. The findings showed that proportions of participants who developed flares during follow-up (24 weeks to four years) ranged from 17% to 35.2%.

Only one study (Senabre 2017)[12] assessed the effect of TDM in people with RA in remission receiving anti-TNFs with extended interval of administration and reported the outcome of proportions of participants who had experienced worsening of clinical activity during the follow-up. The finding from this study (Senabre 2017)[12] showed that 23% participants had experienced worsening of clinical activity at one-year follow-up.

The findings from two prospective uncontrolled cohort studies[58,65] showed that, following the anti-TNF dose tapering strategy (dose reduction), the proportions of participants who achieved persistent remission at follow-up was 87.0% and 92.0%, respectively. The study by Chen and colleagues (2016)[65] had a duration of 24-week follow-up but the study by Rosas and colleagues (2015)[58] did not report the duration of follow-up.

One retrospective uncontrolled cohort study (Paredes and colleagues 2016)[61] evaluated the use of a tapering strategy (dose reduction or discontinuation) of anti-TNF in people with RA with LDA or clinical remission and reported remission rates between pre-visit (baseline) and final visit (four-year follow-up). The results from this retrospective cohort study (Paredes and colleagues 2016)[61] showed that, in comparison with the pre-visit (baseline) remission rate of 77%, 50% of participants maintained remission at final visit after four-year follow-up.

Overall, the evidence from four observational studies generally showed that there was a positive effect in achieving persistent remission associated with TDM for optimising anti-TNF therapies (by decreasing dose or treatment frequency) in participants who had achieved remission or LDA. However, given that these studies were judged to be at moderate risk of bias, there were considerable uncertainties on the reliability of these findings.

Table 25: Changes in disease response, relapse and remission

Study	Study design	Population	Sample size	Missing data (at follow-up)	Length of follow up	Outcome measure	Findings
Senabre 2017[12]	Prospective uncontrolled cohort study	Remission	39	3	1 year	Worsening of clinical activity	23% (9/39)
Chen 2016[65]	Prospective uncontrolled cohort study	Remission/LDA	64	0	24 weeks	Persistent remission	92% (23/25)
						Persistent LDA	62% (24/39)
						Remission turned LDA	0.08 (2/25)
						Disease flare	23% (15/64)
Inciarte-Mundo 2016[63]	Prospective uncontrolled cohort study	Remission/LDA	47	0	1 year	Disease flare	17% (8/47)
Rosas 2015[58]	Prospective uncontrolled cohort study	Patients on remission	45	-	NR	Patients with modified dosing frequency maintaining clinical remission	87%[a]
Paredes 2015[60]	Retrospective uncontrolled cohort study	Remission/LDA	54	-	2 years	Developed flares during follow-up	35.2% (19/54)
Paredes 2016[61]	Retrospective uncontrolled cohort study	Remission/LDA	52	2	4 years	Pre-visit remission/LDA;	77%(40/52)/33%(12/52)
						Final visit remission/LDA/flare	50%(26/52)/27%(14/52/23%(12/52)

Key: LDA: low disease activity; NR: not reported

Notes:

[a] Only proportion reported without actual number

2.3.3.2.2 Change in disease activity

Two observational studies[59,60] evaluated the effect of TDM on change in disease activities at the duration of follow-up of two to seven years. Both studies recruited people with RA who had achieved remission or LDA. The study by Paredes and colleagues (2015)[60] recruited people in remission or LDA for at least six months but the study by Pascual-Salcedo and colleagues (2013)[59] did not report relevant information. The sample size of included studies ranged from 43 to 54. Table 26 presents the results of changes in disease activity.

The study by Pascual-Salcedo and colleagues (2013)[59] had a historical control (i.e. the first period where TDM was not introduced). The findings showed that the mean DAS28 score of participants was 2.51 (standard deviation (SD) 0.85) during historical control period. Compared with the historical control, there was a non-significant reduction in the mean DAS28 score (mean 2.31, SD 0.52) at seven-year follow-up during the second period where TDM was introduced (p=0.061).

The retrospective uncontrolled cohort study by Paredes and colleagues (2015)[60] assessed the outcome measure of DAS28 score by anti-TNF received at pre-visit (baseline) and post-visit with two-year follow-up. The findings showed that, for participants receiving ADL, the mean DAS28 scores were 2.13 (SD 0.12) at pre-visit and 2.42 (SD 0.18) at post-visit (p=0.064). For participants receiving IFX, the mean DAS28 scores were 2.32 (SD 0.11) at pre-visit and 2.19 (SD 0.18) at post-visit (p=0.799). For participants receiving ETN, the mean DAS28 scores were 2.36 (SD 0.12) at pre-visit and 2.93 (SD 0.20) at post-visit (p=0.056). The results indicated that TDM was associated with a non-significant reduction in mean DAS28 score at post-visit after two-year follow-up compared with pre-visit in participants receiving IFX therapies, but non-signficant increases in mean DAS28 scores at two-year follow-up were observed in participants receiving ADL and ETN.

Overall, the finding from the historically controlled study[59] showed that TDM was associated with a non-significant reduction in mean DAS28 scores at seven-year follow-up compared with the historical control. However, mixed results were observed in the retrospective uncontrolled cohort study by Paredes and colleagues (2015).[60] Given the inconsistency in the results, there was uncertainty concerning the impact of TDM on participants' disease activity. It should be noted that the quality of data was judged to be at moderate risk of bias, which compromises the reliability of the findings.

Table 26: Change in disease activity

Study	Study design	Populations (e.g. remission)	Sample size	Missing data (at follow-up)	Length of follow-up	Outcome measure	Findings
Pascual-Salcedo 2013[59]	Historically controlled study	Remission/ LDA	43	NR	7 years	Mean DAS28 score	1st period: 2.51 (SD 0.85) 2nd period*: 2.31(SD 0.52), p=0.061
Paredes 2015[60]	Retrospective uncontrolled cohort study	Remission/ LDA	54	0	2 years	Mean DAS28 score by anti-TNFα received (pre-visit and post-visit, respectively)	ADL: 2.13 (SD 0.12), 2.42 (SD 0.18) (p=0.064) IFX: 2.32 (SD 0.11), 2.19 (SD 0.18) (p=0.799) ETN: 2.36 (SD 0.12), 2.93 (SD 0.20) (p=0.056)

Key: ADL: adalimumab; DAS28: disease activity score in 28 joints; ETN: etanercept IFX: infliximab; LDA: low disease activity; TNF: tumour necrosis factor-alpha

Notes:

*Therapeutic drug monitoring was introduced in the second period

2.3.3.2.3 Change in direction and magnitude of therapeutic dose

Three observational studies[58,59,61] evaluated the outcome of change in direction and magnitude of therapeutic dose in people with RA who had achieved remission or LDA. Both the study by Rosas and colleagues (2015) and the study by Paredes and colleagues (2016) recruited participants who had achieved remission or LDA for at least six months, but the study by Pascual-Salcedo and colleagues (2013) did not report relevant information. The sample size of included studies ranged from 43 to 52. Table 27 presents the results of change in direction and magnitude of therapeutic dose. It should be noted that the results from the study by Pascual-Salcedo et al (2013)[59] on the change of therapeutic dose were presented for the mixed population (including 43 people with RA and 45 people with PsA). Therefore, there was limited generalisability of findings from this mixed population to the target RA population.

The findings from the study by Pascual-Salcedo and colleagues (2013)[59] demonstrated that, compared with the historical control (i.e. the first period where TDM was not used), there were statistically significant reductions in the weekly mean dose per participant by each drug during the second period following the introduction of TDM. For participants receiving IFX, a statistically significant reduction in the weekly mean dose per participant during the second period was observed (mean 0.42 mg/kg/week, SD 0.12), compared with the first period (mean 0.51 mg/kg/week, SD 0.14) ($p<0.001$). For participants receiving ADL, a statistically significant reduction in the weekly mean dose per participant during the second period was also observed (mean 15.52 mg/week, SD 4.81) for the second period, compared with the first period (mean 19.19 mg/week, SD 3.72) ($p<0.001$). Similarly, for participants receiving ETN, there was a statistically significant reduction in the weekly mean dose per participant during the second period (mean 35.04 mg/kg/week, SD13.37) for the second period, compared with the first period (mean 42.09 mg/kg/week, SD 13.25) ($p=0.009$).

The findings from the study by Pascual-Salcedo and colleagues (2013)[59] further showed that, compared with the historical control, there was a statistically significant increase in the mean interval of administration for each drug during the second period where TDM was implemented. For participants receiving IFX, a significantly increased mean interval of administration was observed during the second period (mean 9.7 weeks, SD1.44), compared with the first period (mean 8.52 weeks, SD 1.43) ($p<0.001$). For participants receiving ADL, a significantly increased mean interval of administration was also observed during the second period (mean 2.95 weeks, SD1.58), compared with the first period (mean 2.19 weeks, SD 0.58) ($p=0.007$). Likewise, for participants receiving ETN, a significantly increased mean interval of administration was observed during the second period (mean

1.61 weeks, SD 0.91), compared with the first period (mean 1.09 weeks, SD 0.27) (p=0.004).

Only one prospective uncontrolled cohort study by Rosas and colleagues (2015)[58] assessed the impact of TDM on the total dose of anti-TNFs being avoided. The results demonstrated that the total number of doses avoided was 548 for ETN and 260 for ADL compared with expected dosing schedule, respectively. The study did not report further details of units for doses (e.g. milligrams). This led to cost saving associated with TDM. However, this study did not report the duration of follow-up.

One retrospective uncontrolled cohort study by Paredes (and colleagues (2016)[61] assessed the mean drug levels between the pre-visit (baseline) and post-visit (follow-up) for each anti-TNF (ADL, ETN, and IFX) at the duration of four-year follow-up. The results showed that, compared with those at pre-visit, there were statistically significant reductions in the mean drug levels at post-visit for each anti-TNF being evaluated. For participants receiving ADL, a statistically significant reduction in the mean drug level was observed at post-visit (mean 1507.2, SD 322.7), compared with the pre-visit (mean 5251.9, SD 1205.9) (p=0.001). The unit of measurement was not provided. For participants receiving ETN, a statistically significant reduction in the mean drug level was also observed at post-visit (mean 1,114.9, SD 283.2), compared with the pre-visit (mean 2,735.2, SD 347.4) (p=0.002). Again, for participants receiving IFX, there was a statistically significant reduction in the mean drug level at post-visit (mean 650.4, SD 290.1) compared with the pre-visit (mean 2358.4, SD 728.5) (p=0.008).

Overall, the limited data from three observational studies showed that TDM for optimisation of anti-TNF therapies was associated with reductions in therapeutic dose of anti-TNFs in people with RA who had achieved remission or LDA. This would be expected to lead to cost saving associated with TDM. However, the reliability of findings may be compromised by the poor quality of data being identified.

Table 27: Changes in number, direction and magnitude of dose

Study	Study design	Population (e.g. remission)	Sample size	Missing data (at follow-up)	Length of follow-up	Outcome measure	Findings
Pascual-Salcedo 2013[59]	Historically controlled study	Remission/ LDA	43	NR	7 years	Weekly mean dose per person by drug (1st period vs. 2nd period)*	IFX (mg/kg/week): 0.51 (SD 0.14), 0.42 (SD 0.12) (p<0.001) ADL (mg/week): 19.19 (SD 3.72), 15.52 (SD 4.81) (p<0.001) ETN (mg/week): 42.09 (SD 13.25), 35.04 (SD 13.37) (p=0.009)
						Mean interval of administration by drug (weeks) (1st period vs. 2nd period)*	IFX: 8.52 (SD 1.43), 9.7 (SD 1.44) (p<0.001) ADL: 2.19 (SD 0.58), 2.95 (SD 1.58) (p=0.007) ETN: 1.09 (SD 0.27), 1.61 (SD 0.91) (p=0.004)
Rosas 2015[58]	Prospective uncontrolled cohort study	Remission	45	0	NR	Total number of doses avoided	ETNt: 548** ADL: 260**
Paredes 2016[61]	Retrospective uncontrolled cohort study	Remission/ LDA	52	2	4 years	Mean drug levels*** (pre-visit vs. post-visit)	ADL: 5251.9 (SD 1205.9), 1507.2 (SD 322.7) (p = 0.001) ETN: 2735.2 (SD 347.4), 1114.9 (SD 283.2) (p = 0.002) IFX: 2358.4 (SD 728.5), 650.4 (SD 290.1) (p = 0.008)

Key: ADL: adalimumab, ETN: etanercept, IFX: infliximab; LDA: low disease activity; NR: not reported; SD: standard deviation; vs.: versus

Notes:

*These results were from a mixed population of 43 people with RA and 45 people with PsA

** mg equivalent not reported

*** Unit of measurement not provided

2.3.3.2.4 Discontinuation of ineffective therapy

Only one prospective uncontrolled cohort study by Lopez-Casla and colleagues (2013)[62] evaluated the effect of TDM on treatment decision making and reported the outcome of treatment discontinuation. This study assessed whether TDM for optimising anti-TNF therapies (increasing the dose of IFX) was an effective therapeutic strategy in 36 people with RA who had developed clinical inefficacy (including participants who had experienced a primary non-response or a secondary non-response). The authors did not state the definitions of primary and secondary non-responders.

Table 28 presents the results of discontinuation of ineffective therapy. This study (Lopez-Casla and colleagues 2013)[62] reported that 76.5% of participants discontinued their anti-TNF therapies (IFX) in the one-year follow-up period due to ineffectiveness of anti-TNF therapies. There were no missing data during the period of one-year follow-up in this study. The study concluded that increasing the dose of IFX did not seem to be an effective option in people with RA who had developed clinical inefficacy.

2.3.3.2.5 Adverse effects of treatment

There was only one prospective uncontrolled cohort study (Senabre Gallego and colleagues 2017)[12] reporting adverse effects (AEs) of anti-TNF therapies such as infections. This study recruited 39 people with RA who had achieved remission. Other observational studies did not report AEs of anti-TNF therapies.

The results of AEs of treatment (Table 29) showed that one participant (0.03%) had septic arthritis (serious infectious arthritis) associated with anti-TNF therapies (ADL or ETN) during the one-year follow-up period. There were three participants who were lost to follow-up during the one-year follow-up period.

Table 28: Discontinuation of ineffective therapy

Study (First author / year)	Study design	Population (e.g. remission)	Sample size	Missing data (at follow-up)	Length of follow-up (years)	Outcome measurement	Findings
Lopez-Casla 2013[58]	Prospective uncontrolled cohort study	Primary and secondary non-responders	36	0	1	Treatment discontinuation	76.5%(26/36)[a]

Note:

[a]As reported by the study authors. But the study authors reported that a total of 26 participants discontinued the IFX therapies.

Table 29: Adverse effects of treatment

Study	Study design	Population (e.g. remission)	Sample size	Missing data (at follow-up)	Length of follow-up (years)	Outcome measure	Findings
Senabre Gallego 2017[12]	Prospective uncontrolled cohort study	Remission	39	3	1	Septic arthritis	0.03% (1/39)

2.3.4 Discussion

This systematic review has identified eight studies (reported in 11 publications)[11,12,15,58-65] that evaluated the effect of TDM on clinical outcomes in people with RA who had achieved remission or LDA, or in those people who had experienced a primary non-response or a secondary non-response. Three articles[11,15,64] reported the same non-randomised controlled trial (the INGEBIO study). The remaining studies were observational studies evaluating the impact of TDM.

Most studies recruited people with RA who had achieved remission or LDA. Only one study (Lopez-Casla and colleagues 2013)[62] recruited people with RA who had experienced a primary non-response or a secondary non-response. The majority of included studies used Promonitor ELISA kits for monitoring drug levels and/or anti-drug antibody levels. Three studies in four sources[59-62] used Sanquin ELISA kits to measure drug levels and/or anti-drug antibody levels. It was unclear whether these tests were performed at the centralised testing service. The included studies measured drug levels and/or anti-drug antibody levels in participants who were being treated with ADL, ETN and/or IFX. No studies were identified in participants treated with certolizumab pegol or golimumab. No studies were identified evaluating eligible ELISA kits including IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits.

Comparative controlled evidence

Three abstracts[11,15,64] were identified reporting the same non-randomised controlled trial (the INGEBIO study), which focused on the population who had achieved treatment target (remission or LDA). In this trial, ADL and anti-ADAb levels were measured using Promonitor-ADL and Promonitor-ANTI-ADL (Grifols-Progenika). This trial recruited a mixed population of 169 participants including a cohort of 63 people with RA. The results of the total mixed population were reported in the review as the authors were not able to provide the results for the subgroup of people with RA.

The findings from this non-randomised controlled trial (the INGEBIO study) showed that there was a non-significant reduction in risk of flare in the intervention group compared with the control group. In particular, participants' HRQoL outcomes were higher in the intervention group at all visits compared with the control group, with statistically significant results being observed at two visits. However, as the quality of this trial was judged to be at serious risk of bias, the results should be interpreted with caution. Ideally, randomising participants is required to minimise the risk of bias for the study findings.

Evidence from observational studies

Seven observational studies (reported in eight publications) were identified evaluating the effect of TDM on clinical outcomes in people with RA who had achieved remission or LDA, or in those who had experienced a primary non-response or a secondary non-response. One observational study[59] had a historical control while other studies were single-arm trials with no comparator.

Change in disease response

Five observational studies (reported in six articles) (Chen and colleagues 2016; Inciarte-Mundo and colleagues 2016; Paredes and colleagues 2015; Paredes and colleagues 2016; Chen and colleagues 2016; Rosas and colleagues 2015),[58,60,61,63,65] assessed changes in disease response in people with RA, with sample sizes ranging from 36 to 64. The duration of follow-up of included studies ranged from 24 weeks to four years. All studies focused on people with RA who had achieved treatment target (remission or LDA).

Overall, the evidence from these observational studies generally showed that there was a positive effect in achieving persistent remission associated with TDM for optimisation of anti-TNF therapies (by decreasing dose or treatment frequency) in people with RA who had achieved remission or LDA. However, given that these studies were judged to be at moderate risk of bias, there were considerable uncertainties associated with the reliability of the findings.

Change in disease activity

Two observational studies (Pascual-Salcedo and colleagues 2013; Paredes and colleagues 2015),[59,60] evaluated the effect of TDM on change in disease activities at duration of follow-up of two to seven years, with sample sizes ranging from 43 to 54. Both studies focused on people who had achieved remission or LDA. Overall, the finding from the historically controlled study (Pascual-Salcedo and colleagues 2013) showed that TDM was associated with a non-significant reduction in mean DAS28 scores at seven-year follow-up compared with the historical control (where TDM was not introduced). However, mixed results were found in the retrospective uncontrolled cohort study by Paredes and colleagues (2015).(52) Given the inconsistency of results, there was uncertainty on the impact of TDM on participants' disease activities. It should be noted that the quality of data was judged to be at moderate risk of bias, which has compromised the reliability of the findings.

Change in direction and magnitude of therapeutic dose

Three observational studies[58,59,61] evaluated the outcome of changes in direction and magnitude of therapeutic dose in people with RA who had achieved remission or LDA. The sample size of included studies ranged from 43 to 52.

Overall, the limited data from three observational studies showed that TDM for optimising anti-TNF therapies was associated with reductions in therapeutic dose of anti-TNFs in people with RA who had achieved remission or LDA. This would be expected to lead to cost saving associated with TDM. Where statistically significantly results were observed, these results may be clinically significant. However, the reliability of the findings may be compromised by the poor quality of data being identified.

2.3.4.1 Reliability of the findings

The non-randomised controlled study[11,15,64] was judged to be at serious risk of bias (see Section 2.3.2.4). In this trial, there was an issue of baseline imbalance in disease severity between the intervention and control groups. Furthermore, there was a lack of adjusting for this variable in the analysis of clinical outcomes. There were higher attrition rates for some outcomes, which can lead to attrition bias. These deficiencies resulted in serious risk of bias associated with the findings. Therefore, the results should be interpreted with caution.

In terms of observational studies, the historically controlled study and all six single arm studies were judged to be at moderate risk of bias (see Section 2.3.2.4). The study by Pascual-Salcedo (2013)[59] was judged to be at moderate risk of bias because there was non-contemporaneous control bias due to the use of a historical control. It should be noted that the same group of participants were assessed during the first period (the historical control where TDM was not introduced) and the second period (where TDM was implemented). However, most observational studies had a small sample size without a control group. Therefore, the overall poor quality of included studies compromises the reliability of the findings.

2.3.4.2 Generalisability of the findings

Given that most studies were conducted in Spain, the findings from these studies may have limited generalisability to the UK setting due to variations in clinical practice and health policies between different countries. Furthermore, the findings from the non-randomised controlled trial (the INGEBIO study) and the results of changes in therapeutic dose from the study by Pascual-Salcedo and colleagues (2013)[59] were presented for a mixed population. Therefore, there was limited generalisability of findings from the mixed population (including RA, PsA and/or ankylosing spondylitis) to the target RA population.

2.3.4.3 Implications for future research

One ongoing Norwegian multicentre RCT was identified (the NOR-DRUM Study)[66] that evaluates the effect of TDM in people with RA in remission compared with standard care (see Section 2.3.2.3). This ongoing trial will provide further useful data on the impact of TDM in the target population.

Further controlled trials with a large sample size (especially RCTs) are required to assess the impact of using Promonitor ELISA tests for monitoring anti-TNF therapies in people with RA who had achieved remission or LDA.

No studies were identified that assessed other eligible ELISA kits including IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits.

Therefore, future large RCTs are required to assess the impact of using those ELISA tests for monitoring anti-TNF therapies in people with RA who had achieved remission or LDA. More robust evidence is also needed to evaluate the impact of using Sanquin tests for monitoring anti-TNF therapies in this population

Future RCTs are warranted to evaluate the clinical effectiveness of using ELISA tests for monitoring anti-TNF therapies in people with RA who had experienced a primary non-response or a secondary non-response.

There were no studies identified for patients who were being treated with certolizumab pegol and golimumab. Future RCTs are required to assess the clinical effectiveness of using ELISA tests for monitoring such anti-TNF therapies in the target populations.

2.3.4.4 Conclusions

In relation to clinical effectiveness, limited data were identified evaluating TDM in the target populations. One non-randomised trial compared TDM with standard care (the INGEBIO study) had serious limitations in relation to the NICE scope: only one-third of the participants had RA, many of the analyses were not by intention-to-treat, follow-up was for only 18 months, there was no explicit algorithm for guiding clinicians in how the results of testing should change treatment (e.g. tapering), and the study was only reported in three abstracts. In addition, seven observational studies (reported in eight publications) were also identified but were of minimal value in informing whether ELISA test-based monitoring is clinically effective or not.

3 Systematic review of existing cost-effectiveness evidence

3.1 Objectives

The objectives of this systematic review of economic evaluations were as follows:

- To gain insights into the key drivers of cost-effectiveness of TNF testing.
- To get an overview of the alternative modelling approaches that have been adopted to evaluate the use of therapeutic drug monitoring in people with RA.
- To provide a summary of the findings of previous relevant cost-utility, cost-effectiveness, and cost-benefit studies.

3.2 Methods

3.2.1 Identification of studies

The following bibliographic databases were searched:

- MEDLINE (Ovid)
- MEDLINE In-Process & Other Non-Indexed Citations (Ovid)
- Embase (Ovid)
- Web of Science (Thomson Reuters)
- NHS EED and HTA (The Cochrane Library)
- EconLit (EBSCO)

In addition, searches were carried out on the following websites: HUD (ScHARR), HERC (Oxford), EQ-5D (EuroQol), CEA Registry and ISPOR.

The searches were developed and run by an information specialist (SR) in July 2018 and updated in November 2018. They comprised terms for RA and terms for anti-TNF inhibitors and terms for enzyme linked immunosorbent assay (ELISA) testing. Search filters were used to limit the searches to cost-effectiveness studies. No date or language limits were used.

Separate searches were also carried out for appropriate health utilities and costs, using a variety of search terms and filters. These searches were carried out in several iterations to look for different aspects of costs and health utilities for RA and ELISA tests as needed.

The full search strategies for each database, for cost effectiveness and one example iteration of the utilities searches are provided in Appendix 1. The database search results were exported to, and deduplicated using Endnote (X7). Deduplication was also performed using manual checking.

Screening was done independently by two reviewers. Disagreements between reviewers were resolved by consensus. All references considered for inclusion by either reviewer at the title and abstract stage were included for full-text screening.

3.2.2 Eligibility criteria

Eligible studies for inclusion to the systematic review were selected according to inclusion and exclusion criteria outlined in a PICO template. The inclusion criteria for population, interventions and comparator were as described in Section 2.2.1.1 and Section 2.2.1.2. The following types of economic evaluations were included: cost-utility analyses, cost-effectiveness, cost-benefit, cost-consequence and cost-minimisation analyses. Systematic reviews of economic studies were also considered.

3.2.3 Data extraction

Study characteristics and results were extracted and summarised by one reviewer (MR). The evidence was assessed using narrative synthesis supported by summary data extraction tables.

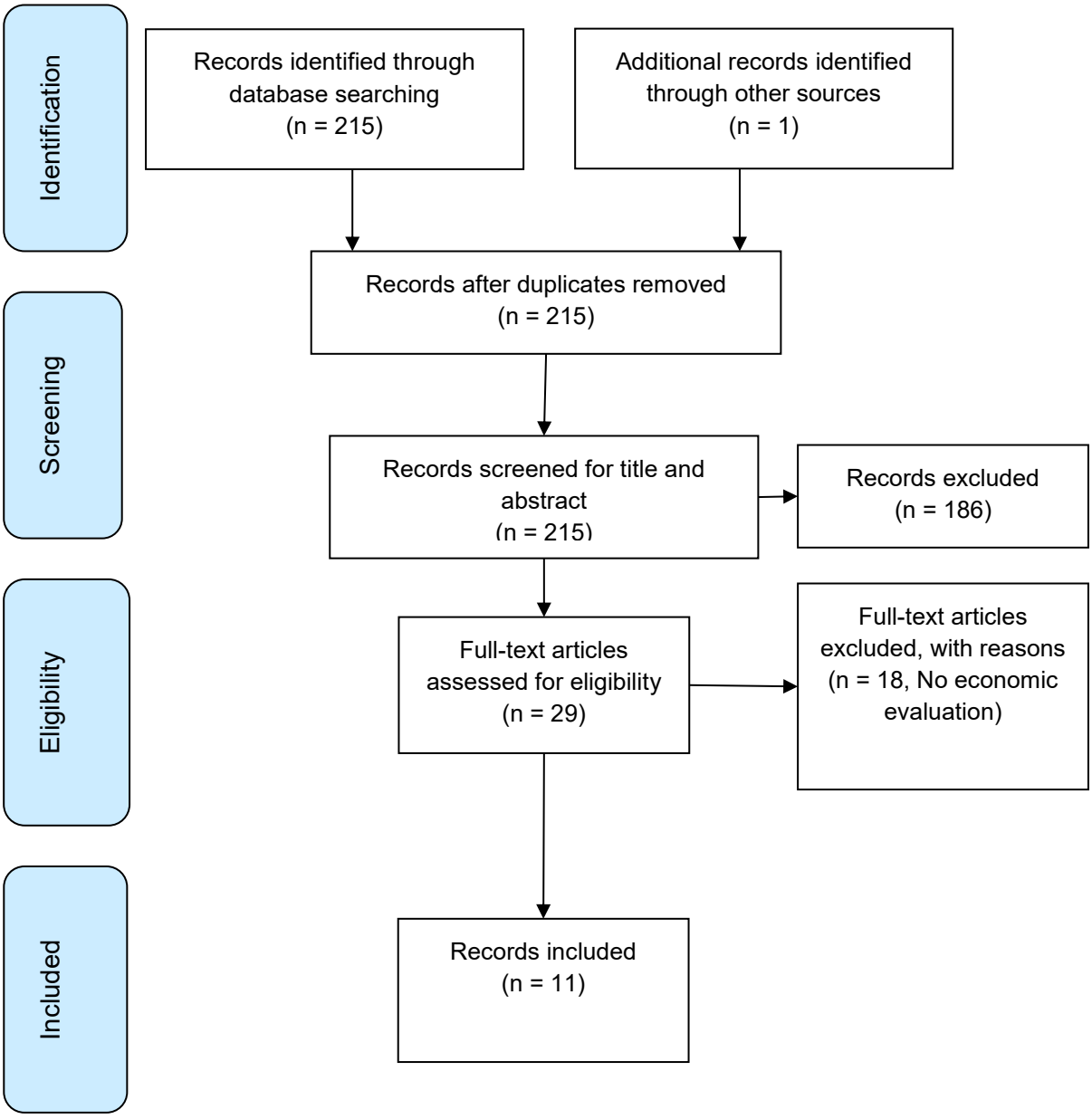
3.2.4 Critical appraisal

The quality of the selected studies was evaluated by one reviewer according to the Consensus Health Economic Criteria (CHEC).[68] Studies based on decision models were further quality-assessed using the checklist developed by Philips and colleagues (2006).[69]

3.3 Results

Figure 3 shows a PRISMA flow diagram for the systematic review.[55] After deduplication, 214 records were identified. All records were screened on title and abstract, and 29 citations were screened at full text. In addition to the records identified from searches of electronic sources, we were made aware of a PhD thesis (unpublished, personal communication) that met inclusion criteria for the cost-effectiveness systematic review. Overall, 11 sources met the inclusion criteria for the systematic review.

Figure 3: PRISMA flow diagram: a description of study inclusion process for the cost-effectiveness systematic review



Five studies (reported in 11 publications) were eligible for inclusion (Table 30): three studies were model-based economic evaluations. Of these, two were reported only in abstract format: one was a non-randomised controlled trial (INGEBIO), and one was an observational study (Pascual-Salcedo and colleagues 2013). The authors of the abstracts were contacted and provided two poster presentations reporting outcomes of the INGEBIO study. These sources are not included in the PRISMA diagram.

Table 30: Characteristics of included studies

Author	Type of reference	Type of study	Sources
INGEBIO	Abstract	Non-randomised controlled trial	Arango and colleagues (2017),[15] Ucar and colleagues (2017),[11] and Gorostiza and colleagues (2016)[64]
Krieckaert and colleagues	Full text	Model	Krieckaert and colleagues (2012)[70,71] Krieckaert and colleagues (2013)[72,73] Krieckaert and colleagues (2015)[1]
Pascual-Salcedo and colleagues.	Abstract	Observational	Pascual-Salcedo and colleagues (2013)[59]
Laine, J. and colleagues	Full text	Model	Laine and colleagues (2016)[2]
Gavan, S.	Dissertation	Model	Personal communication

Characteristics of the included studies are given in Table 31 and Table 32.

Table 31: Observational cost-effectiveness studies of therapeutic drug monitoring tests in people with rheumatoid arthritis

Study	Population	Setting	Test	TNF-α inhibitor	Study design	N	Time-frame	Outcome	Cost measures	Results	Comments
Ucar 2017 (INGEBIO)	People with RA,PsA and AS, treated with ADL who remained clinically stable for at least 6 months	Clinic, Spain	Trough ADL and ADAb measured by Promonitor-ADA and Promonitor anti-adalimumab (Progenika) at 8 time points	ADL	Non-randomised controlled trial	109 participants in IG and 60 in CG, of which 30 and 33 people with RA, respectively	18 months	DAS28, BASDAI, BASFI and HAQ-DI, days with active disease	Average cost of ADL per patient-year	Mean QALY were 1.145 and 1.076 during follow-up period per person in IG and CG, respectively; the average cost of Humira® (ADL) per patient-year was 10,664.54€ vs 9,856.45€ (-808.08€, 8% savings) in the CG and IG, respectively (the results reported for the mixed population)	Data is reported for all participants and is not reported by disease subgroup. People with rheumatic diseaase have better quality of life, lower risk of flares and incur lower treatment costs if management is complemented with ELISA testing.

Study	Population	Setting	Test	TNF-α inhibitor	Study design	N	Time-frame	Outcome	Cost measures	Results	Comments
Pascual-Salcedo 2013	People with RA and SpA in remission or LDA under treatment by IFX, ADA and ETN	University Hospital, Spain	Drug levels by capture ELISA	ADL, IFX, ETN	Historical controlled cohort	43 participants with RA		DAS28	Monthly amount of spared drug per person	Decrease in drug use: €91.62 per person for IFX (70 kg of mean weight), €324 per person for ADL, €257 per person for ETN	Data is reported for all participants and is not reported by subgroup. QALYs are not reported

Key: ADAb: anti adalimumab antibody; ADL: adalimumab; AS: ankylosing arthritis; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; CG: control group; DAS28: Disease activity score in 28 joints; ELISA: enzyme linked immunosorbent assay; ETN; etanercept; HAQ-DI: health assessment questionnaire disability index; IFX: infliximab; IG: intervention group; LDA = low disease activity; PsA: psoriatic arthritis; RA: rheumatoid arthritis; SpA: spondyloarthritis

Table 32: Modelling studies of using therapeutic drug monitoring in people with rheumatoid arthritis

Study	Population	Perspective	Setting	Test	TNF-α inhibitor	Model structure	Time-frame	Effectiveness and costs parameters	Type of study	Comments
Krieckaert 2015	Cohort of 272 people with RA treated with ADL for 3 years [1], and data on direct medical costs and HRQoL from the URAC study group (N=1,034)	Societal, Healthcare	Clinic, Netherlands	Sanquin, DL	ADL	Markov	3-year horizon with 3-month cycles	Direct medical and productivity costs, utilities	Cost-utility	The authors compared a cohort of monitored by ELISA testing, with a cohort from the URAC study at 28 weeks after starting ADL. Markov states were defined by DAS28 categorisation. Result: substantial reduction in costs of medication, small change in efficacy of treatment for ELISA testing.
Laine 2016	Cohort of people treated with ADL (N=486) and IFX (N=1,137)[2]	Healthcare	Clinic, Finland	Sanquin, Promonitor, immunoassay,DL and Ab	ADL, IFX	Markov	3-year horizon with 6-month cycles	Decreasing proportion of people on non-optimal treatment; costs of test and non-optimal treatment	Cost-effectiveness	Economic impact of clinical decision-making was modelled in a short-term (3–6 months) scenario with 100 hypothetical patients for ADL and IFX. ELISA testing was performed in non-responders. Result: using ELISA test was cost-saving.

Study	Population	Perspective	Setting	Test	TNF-α inhibitor	Model structure	Time-frame	Effectiveness and costs parameters	Type of study	Comments
Gavan 2017	People with RA in England (BSRBR-RA)	NHS and PSS	England	ELISA tests, no specific ELISA test stated	ADL	DES	Lifetime	Costs of treatment, hospitalisation and testing	Cost-utility	ELISA monitoring was investigated for use during response and in remission for dose adjustment. Result: ELISA testing is not likely to be cost-effective.

Key**:** Ab: antibody; ADL: adalimumab; IFX: infliximab; DES: discrete-event simulation; DL: drug level; ELISA: enzyme linked immunosorbent assay; NHS: National Health Service; PSS: personal social services; RA: rheumatoid arthritis; URAC: Utrecht Rheumatoid Arthritis Cohort

Notes:

[1] This was a prospective observational cohort study of 272 people with RA treated with ADL therapy at the Department of Rheumatology, Jan van Breemen Institute, Amsterdam, Netherlands. All participants fulfilled the ACR 1987 revised criteria for RA and had active disease indicated by a DAS28 of at least 3.2, despite earlier treatment with two disease-modifying antirheumatic drugs (DMARDs) including methotrexate at 25 mg weekly or at the maximal tolerable dosage, according to the Dutch consensus statement on the initiation and continuation of TNF–blocking therapy in RA.

[2] The data were obtained from the clinical sample registry of United Medix Laboratories Ltd in Helsinki, Finland. All the samples included in the database were sent to the laboratory on a clinical basis.

3.3.1 Non-model based studies

The two studies, Ucar and colleagues (2017) and Pascual-Salcedo and colleagues (2013), were reported as abstracts. Ucar and colleagues (2017) investigated the impact of monitoring of ADL drug level and anti-drug antibody level in people with RA, psoriatic arthritis (PsA) and ankylosing spondylitis (AS) on annual direct costs to the Health System and health outcomes compared to conventional practice in Spain. The economic analysis reported in Ucar and colleagues (2017) was based on clinical outcomes from a pragmatic, non-randomised, non-inferiority study. Trough ADL level and anti-drug antibody level were measured with Promonitor-ADL and Promonitor-ANTI-ADL. Physicians were not obliged to adhere to any therapeutic algorithm when making treatment decisions for participants in the intervention group. In the control group, treatment decisions were based on clinical judgment only. 169 people were recruited, of whom 63 (37.3%) had RA (30 people with RA in the intervention group and 33 people with RA in the control group). Ucar and colleagues (2017) reported the result for all participants and did not report results not by subgroup (disease categories). It is therefore difficult to generalise the results to people with RA. The authors reported that people with RA from the intervention group had better quality of life, lower risk of flares and lower treatment costs when compared to the control group. The average cost of Humira® (ADL) per patient-year was €10,664.54 vs. €9,856.45 (-€808.08, 8% savings) in the control and intervention arms, respectively; the results were reported for the total (mixed) population. The average annual cost of ADL treatment per participant was reported.

The Pascual-Salcedo and colleagues (2013) study aimed to compare the clinical and economic impact of TDM, based on serum trough drug levels, in people with RA and SpA in remission or with low disease activity (LDA). This was an observational study of routine clinical practice. The study included a total of 88 participants (43 RA and 45 SpA), treated with three TNF inhibitors (31 withIFX, 29 with ADL and 28 with ETN). Participants were followed for seven years (2006-2012). Drug levels were measured using ELISA test. No further information on the test was given in the abstract. For each participant two time periods were examined, one before and the other during TNF drug monitoring, 2006-2009 and 2010-2012, respectively. All participants in this study had stable clinical activity in both time periods. Pascual-Salcedo and colleagues (2013) reported the monthly amount of spared drug as €91.62 per participant treated with IFX (assuming the mean participant weight of 70 kg), €324 per participant on ADL, and €257 per participant on ETN.

3.3.2 Model-based studies

Three model-based economic evaluations were identified in the systematic review (Table 32). All were conducted in Europe (Netherlands, Finland and the UK).

3.3.2.1 Krieckaert 2015

The study reported by Krickaert and colleagues (2015) was a cost-utility study investigating the role of testing ADL drug levels in people with RA. Drug levels were measured using in-house ELISA tests (Sanquin, Amsterdam) in a cohort of 272 ADL-treated people with RA recruited at the Department of Rheumatology, Jan van Breemen Institute, Amsterdam, for three years (Bartelds and colleagues, 2011[74]). These participants were compared with a cohort of 1,034 participants from the Utrecht Rheumatoid Arthritis Cohort (URAC) treated with other treatments based on clinical judgment. The clinical characteristics of these participants are not clearly discernible from the cited references. Participants in the intervention cohort were tested at four, 16, 28, 40 and 52 weeks of treatment and every six months thereafter. However, in the economic analysis, the authors modelled ELISA testing at 28 weeks only (Figure 4). After three years, 76 of a total 272 participants (28%) developed anti-ADL antibodies; 51 of these (67%) during the first 28 weeks of treatment. Over the course of the study, participants with measurable antibody levels were 3.2 times less likely (95% CI for HR 1.8, 7,2) to revert to minimal disease activity and 7.1 times less likely (95% CI for HR 2.1, 23.4) to enter sustained remission, based on DAS28 scores <3.2 and <2.6, respectively. Clinical outcomes from Bartelds and colleagues (2011) are summarised in Table 33.

Table 33: Clinical outcomes from Bartelds and colleagues (2011)

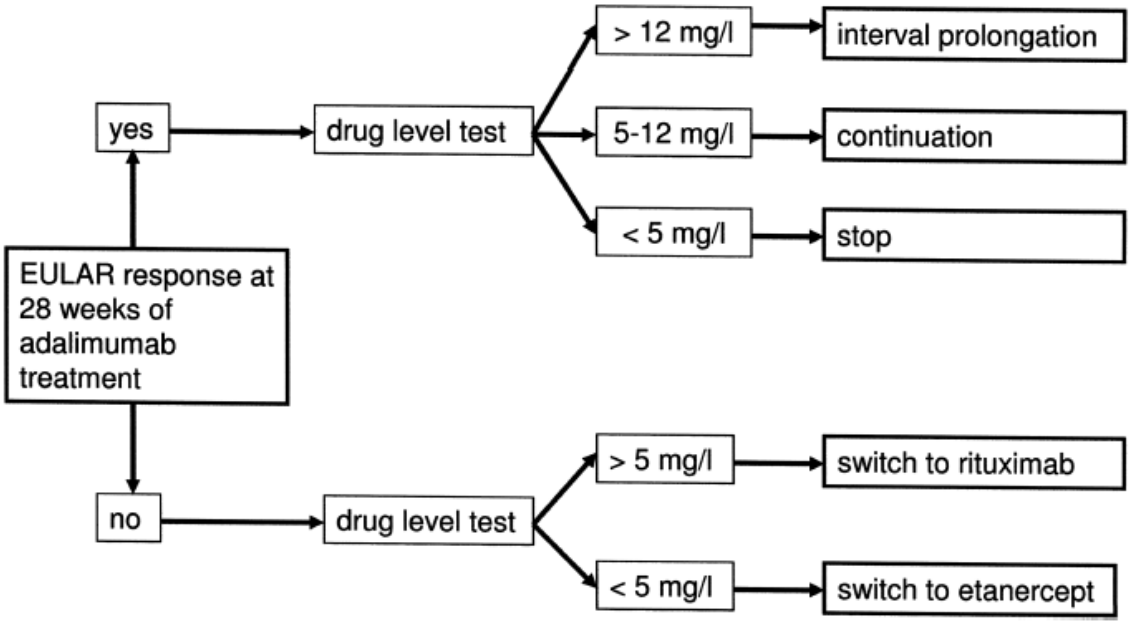
Antibody titre (AU/mL)	Drug Level (median/IQR,mg/L)	Treatment discontinuation (%)	Disease activity (%)	Sustained remission (%)
Undetectable(n=196)	12/9-16	28(14)	Minimal = 95 (48)	67 (34)
13 to 100 (n=45)	5/3-9	29(38)	Minimal = 10 (22)	3 (7)
>100 (n=31)	0/0-3			

Key: IQR: interquartile range
Source: Bartelds and colleagues (2011)[74]

Of note, this study was excluded from the clinical effectiveness systematic review because it did not meet the inclusion criteria for the population (the population was treatment naïve and disease active). As this paper was excluded at the first screening stage (titles and abstracts), it was not described in the list of excluded studies (Appendix 2).

The treatment algorithm used in Krieckaert and colleagues (2015) is shown in Figure 4.

Figure 4: Algorithm for decision following the result of test in 28th week after adalimumab in Krieckaert and colleagues (2015)



Source: Krieckaert and colleagues (2015)[1]

The authors used a Markov model with three month cycles and a time horizon of three years using microsimulation for analysis. The analysis was performed probabilistically. Discounting was applied at 4% for costs and 1.5% for utilities in accordance with Dutch national guidelines. Results were reported from both healthcare and societal perspectives. The Markov model health states were based on categorisation of DAS28 as below:

- remission (DAS28 <2.6)
- LDA (2.6≤DAS28<3.2)
- moderate disease activity (3.2≤DAS28≤5.1)
- high disease activity (DAS28>5.1).

Transition probabilities were estimated using a regression function derived from the URAC cohort outcome data (Bartelds and colleagues, 2011[74]). Costs included direct medical and productivity costs. Utility was calculated based on the EQ-5D classification outcomes recorded in the URAC study.

ELISA testing was cost-saving from both the societal and healthcare perspective (Table 34). The test-based treatment strategy resulted in lower costs (due to the reduction in the treatment cost) and greater quality-adjusted life years (QALYs).

Table 34: Cost-effectiveness results reported in Krickaert and colleagues (2015)

Perspective	Costs		QALYs		ICER
	intervention	control	intervention	control	
Societal	€15,466,869	€18,028,517	591.65	587.81	−€ 646,266
Healthcare provider	€13,607,067	€16,153,357	591.65	587.81	−€ 666,541

Key: ICER, incremental cost effectiveness ratio; QALYs, quality-adjusted life years.
Note: Cost year was not reported.

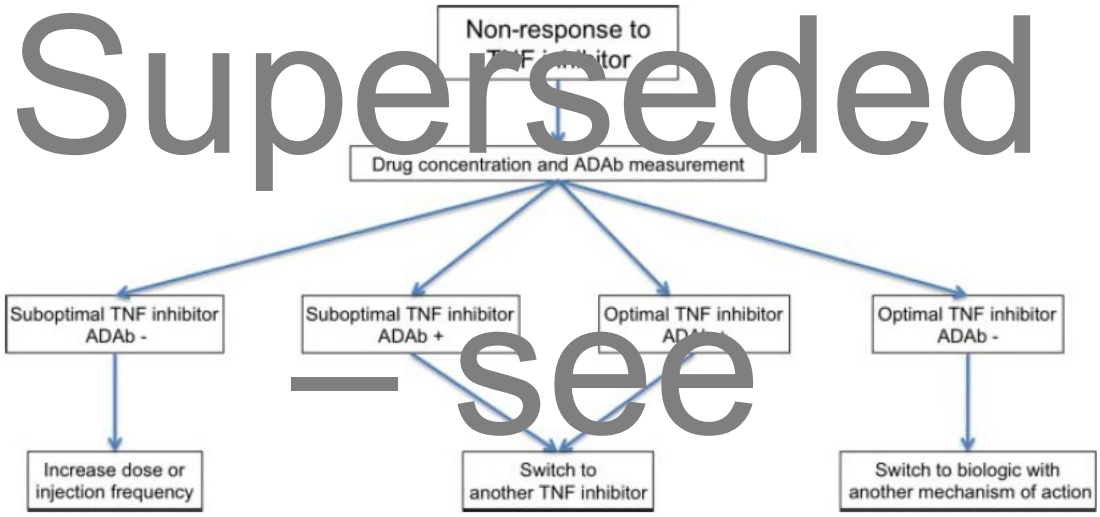
Probabilistic sensitivity analysis around the base case scenario predicted that ELISA testing would dominate usual care in 72% of scenarios. Scenario sensitivity analyses around; e.g. the drug level cut-offs used, or the definitions of a good EULAR response, showed that ELISA testing is generally cost-saving, although some scenarios reported loss of QALYs.

3.3.2.2 Laine and colleagues 2016

Laine and colleagues (2016) conducted a cost-effectiveness study in Finland. The intervention involved assessment of drug and anti-drug antibody levels in people with RA treated with ADL or IFX. The data on drug and anti-drug antibody levels were taken from the clinical sample registry of United Medix Laboratories Ltd in Helsinki, Finland, which included 486 and 1,137 samples from participants on ADL and IFX, respectively. The drug levels were measured using ELISA, while antibody level was assessed using radioimmunoassay. All measurements of antibody and ADL drug levels were outsourced to Sanquin Diagnostic Services (Amsterdam, Netherlands). Approximately half of IFX drug level was measured by the United Medix Laboratories using Promonitor test kit (Progenica, Derio, Spain).

Clinical management decisions based on the test results followed the algorithm proposed by Vincent et al (2013)[75] (Figure 5). Possible treatment decisions included increasing dose, switching to another TNF-α inhibitor or switching to a bDMARD with a different mechanism of action.

Figure 5: Algorithm for interpretation of test results



Source: Vincent 2013

The economic impact of clinical decision-making was modelled in a short-term (three to six months) scenario with 100 hypothetical non-responders. They were compared with a non-testing scenario in which the same participants were managed only by clinical judgment in routine practice. The outcome measures were the changes in the probability of undergoing periods of sub-optimal treatment, and the cost-effectiveness of routine monitoring compared to clinical judgement only. An inappropriate clinical decision was defined to lead to ineffective treatment for at least three to six months. The authors justified this time period based on the typical follow-up visit frequencies of people with RA treated with biologics in Finland (no data sources were provided). This meant that all participants in the control arm experienced a three-month delay in receiving appropriate treatment. This delay was estimated to cost €1,471 for every month, which included the drug cost estimate per month of subcutaneous TNF-α inhibitor (€1,140), travel and lost working and leisure time costs for a laboratory visit (€17.4), costs of the possible standard safety-related laboratory tests (€6.8), travel and lost working and leisure time cost for a follow-up visit to an outpatient specialist clinic (€66.6), and specialist visit (€240.6). Long-term efficacy-related costs were not modelled. The cost of resource use was valued according to the national unit costs inflation adjusted to the year 2013.

The authors proposed a Markov model with six-month cycles and three-year time horizon. Health states were defined as:

- first TNF-α blocker
- second biological (TNF-α blocker or non-TNF drug)
- quitting biologics.

The model predicted that over the three-year period, in the intervention arm, 40% of participants on ADL and 50% of participants on IFX, respectively, will need drug modification. Based on a hypothetical cohort of 100 participants, the cost of testing was estimated to amount maximally to €20,000 (€200×100 participants). Dividing the cost of test by the cost per month of non-optimised treatment then indicates the threshold number of person-months of sub-optimal treatment that correspond with testing being considered cost-effective. Laine and colleagues (2016) reported that the routine measurement of both drug and antibody levels would be cost-saving comparing to the non-testing scenario, assuming that a minimum of 2.5% or 5% of patients are treated non-optimally for six or three months, respectively.

3.3.2.3 Gavan 2017

In the PhD thesis by Gavan (personal communication, 6 August, 2018), the cost-effectiveness of using ELISA testing for monitoring of people with RA treated with ADL was evaluated. Twelve different ELISA test-based strategies were compared against the current practice in England (i.e. no TNF testing) (Table 35).

Table 35: Strategies compared in the Gavan (2017) study

Strategy	Type of testing strategy	Description
Current practice	Not applicable	Usual care for people with RA with no testing of ADAb or drug level
Strategy 1	Monitoring	ADAb and drug level testing every 3 months
Strategy 2	Monitoring	ADAb and drug level testing every 6 months
Strategy 3	Monitoring and dose reduction	ADAb and drug level testing every 3 months, drug level test in remission after 2 years
Strategy 4	Monitoring and dose reduction	ADAb and drug level testing every 3 months, drug level test in remission after 3 years
Strategy 5	Dose reduction	Drug level test in remission after 2 years
Strategy 6	Dose reduction	Drug level test in remission after 3 years
Strategy 7	Monitoring	ADAb testing only every 3 months
Strategy 8	Monitoring	ADAb testing only every 6 months

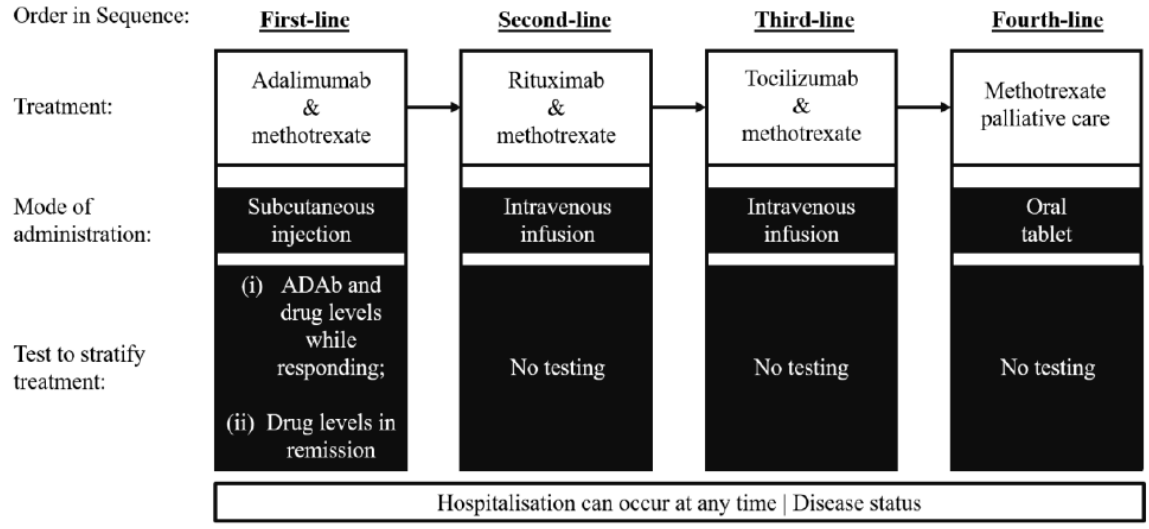
Strategy	Type of testing strategy	Description
Strategy 9	Monitoring and dose reduction	ADAb testing only every 3 months, drug level test in remission after 2 years
Strategy 10	Monitoring and dose reduction	ADAb testing only every 3 months, drug level test in remission after 3 years
Strategy 11	Not applicable	No testing. Just half dose in remission after 2 years
Strategy 12	Not applicable	No testing. Just half dose in remission after 3 years

Key: RA, Rheumatoid arthritis; ADAb, anti-adalimumab antibody.
Notes:
Monitoring was done in responding participants, dose reduction was implemented in patients in remission state.
Source: Gavan (2017)

These strategies were a combination of using monitoring tests during response to the drug and after remission. The author considered a frequency of testing of every three or six months in responders to therapy. For people in remission, testing was considered at two and three years of being in remission.

In Gavan (2017), four lines of treatment were modelled as shown in Figure 6.

Figure 6: Service pathway of RA treatment in England



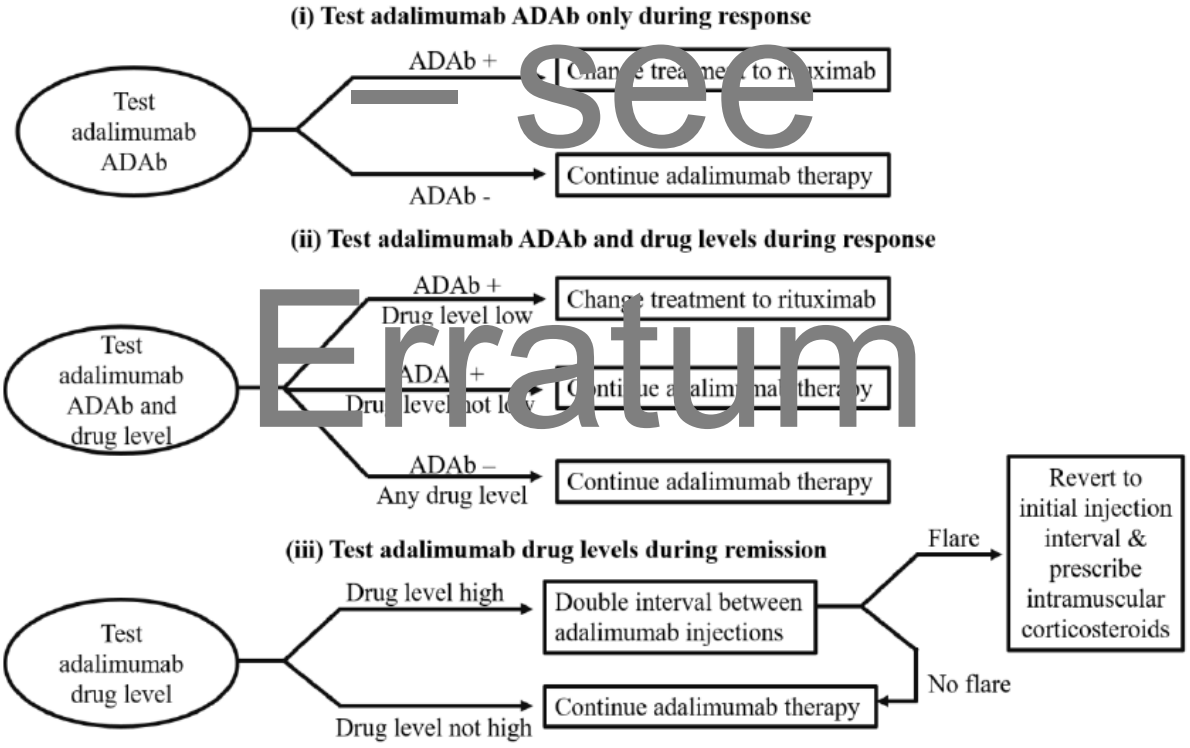
Source: Gavan 2017

A discrete-event simulation (DES) modelling approach was employed. The following competing events were considered: time to death, ADL failure, rituximab failure, tocilizumab failure, time to development of antibodies against ADL, remission, EULAR response and HAQ progression. The model simulated 20,000 hypothetical patients, representative of the

population with RA in England using summary attributes of patients from the British Society for Rheumatology Biologics Register – Rheumatoid Arthritis (BSRBR-RA).

One of the test strategies considered in Gavan (2017)[35] was monitoring of drug and antibody levels in participants responding to treatment in order to avoid the harm associated with secondary non-response. Another possible test strategy was dose adjustment in patients in remission, informed by the results of TNF testing. Figure 7 shows the algorithm used in the Gavan study for management decisions in participant in whom TDM was performed.

Figure 7: Algorithm for test interpretation used in Gavan (2017)



Source: Gavan (2017)[35]

Utilities were calculated based on mapping the HAQ score from the BSRBR-RA using a quadratic mapping algorithm estimated previously for the NICE TA195 by Malottki and colleagues (2011).[14] Costs included the costs of treatment, hospitalisations and testing. Quantities of resource utilisation were derived from published sources (Stevenson and colleagues, 2016;[50] Jani and colleagues, 2016[53]), unit costs were taken from the NHS reference costs 2015-2016 and BNF (accessed 8 April 2016).

Based on the 12 scenarios tested (Table 35), Gavan (2017)[35] concluded that routine use of ADL testing was cost-effective compared to current practice, but was unlikely to be cost-effective relative to dose reduction (without testing) for people in remission (Scenario 11). Overall, two scenarios (7 and 10) showed a negative net monetary benefit compared to

standard care in the analysis. Amongst the set of the remaining ten scenarios, all except three were shown to be dominated or extendedly dominated by another of the set. Finally, of the three remaining strategies the ICERs for strategies 1 (adalimumab antibody [ADAb] and drug level testing every three months) and three (ADAb and drug level testing every three months, drug level test in remission after two years) relative to strategy 11 (no testing, just half dose in remission after two years) were £38,575 and £ 37,043. Since strategy 11 consists merely of dose reduction after two years for people in remission, the analysis of the chosen scenarios therefore suggests that testing may not be cost-effective. Strategy 11 shows an incremental net monetary benefit per patient of £3,896.72 relative to standard care, with an associated mean QALY loss of -0.000121 per patient compared to standard care.[35]

3.4 Quality of identified cost-utility studies

Table 36 shows the results of assessing the included studies against the Consensus Health Economic Criteria (CHEC).[68] Methodological quality of included modelling studies assessed using the Philips checklist[69] is addressed in Table 37.

Table 36: Quality appraisal of cost-utility studies using the CHEC checklist

Item	CHEC-list	Ucar 2017	Pascual-Salcedo 2013	Krieckaert 2015	Laine 2016	Gavan 2017
1	Is the study population clearly described?	Y	Y	Y	N	Y
2	Are competing alternatives clearly described?	Y	Y	Y	Y	Y
3	Is a well-defined research question posed in answerable form?	Y	Y	Y	Y	Y
4	Is the economic study design appropriate to the stated objective?	Y	Y	Y	N	Y
5	Is the chosen time horizon appropriate in order to include relevant costs and consequences?	Y	N	Y	Y	Y
6	Is the actual perspective chosen appropriate?	N	N	Y	Y	Y
7	Are all important and relevant costs for each alternative identified?	N	N	N	N	Y
8	Are all costs measured appropriately in physical units?	N	N	N	N	Y
9	Are costs valued appropriately?	N	N	N	N	Y
10	Are all important and relevant outcomes for each alternative identified?	Y	N	Y	N	Y
11	Are all outcomes measured appropriately?	N	N	Y	Y	Y
12	Are outcomes valued appropriately?	N	N	Y	Y	Y
13	Is an incremental analysis of costs and outcomes of alternatives performed?	N	N	Y	N	Y
14	Are all future costs and outcomes discounted appropriately?	N	N	Y	N	Y
15	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	N	N	Y	N	Y
16	Do the conclusions follow from the data reported?	Y	Y	Y	Y	Y
17	Does the study discuss the generalisability of the results to other settings and patient/client groups?	N	N	N	Y	Y
18	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	N	N	N	N	Y
19	Are ethical and distributional issues discussed appropriately?	N	N	Y	Y	Y

Source: Evers and colleagues (2005)[68]

Table 37: Quality appraisal of cost-utility studies using the checklist developed by Philips and colleagues

	Krieckaert 2015	Laine 2016	Gavan 2017
Structure (S):			
S1: Statement of decision problem/objective	Yes	Yes	Yes
S2: Statement of scope/perspective	Yes	Yes	Yes
S3: Rationale for structure	No	Yes	Yes
S4: Structural assumptions	No	No	Yes
S5: Strategies/comparators	Yes	Yes	Yes
S6: Model type	Yes	Yes	Yes
S7: Time horizon	Yes	Yes	Yes
S8: Disease states/pathways	Yes	No	Yes
S9: Cycle length	Yes	Yes	NA
Data (D):			
D1: Data identification	Yes	Yes	Yes
D2: Pre-model data analysis	No	No	No
D2a: baseline data	Yes	Yes	Yes
D2b: treatment effects	No	No	Yes
D2c: quality-of-life weights (utilities)	Yes	No	Yes
D3: Data incorporation	No	No	No
D4: Assessment of uncertainty			
D4a: methodological	No	No	No
D4b: structural	No	No	No
D4c: heterogeneity	Yes	No	Yes
D4d: parameter	Yes	No	Yes
Consistency (C):			
C1: Internal consistency	Yes	Yes	Yes
C2: External consistency	No	No	No

Key: NA, not applicable
Source: Philips and colleagues (2006)[69]

3.5 Discussion

A systematic literature search performed in July 2018 and updated in November 2018 identified five publications relevant to the decision problem, with two of these available in abstract format only. Furthermore, only two (out of six) TNF testing kits from the NICE scope (Promonitor and Sanquin) and three (out of five) TNF-α inhibitors (ADL, ETN, and IFX) were considered in the selected studies (Table 38).

Table 38: Cost-effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope

		Promonitor	IDKmonitor	LISA-TRACKER	RIDASCREEN	Sanquin*
ADL	*drug*	✓[1]	X	X	X	✓[2]
	antibody	✓[3]	X[4]	X	X	✓[5]
ETN	*drug*	✓[6]	X	X		X
	antibody	X	X	X		X
IFX	*drug*	✓[6]	X[4]	X	X	X
	antibody	X	X	X	X	✓[6]
Golimumab	*drug*	X	X	X		X
	antibody	X	X	X		
Certolizumab pegol	*drug*			X		X
	antibody			X		

Key:

X Indicates availability of a test to measure drug or antibody level in participants treated with the specified TNF-α inhibitor and that no studies have been identified in the clinical-effectiveness systematic review, reporting on using therapeutic drug monitoring for the specified test kit and TNF-α inhibitor.

✓ Indicates availability of a test to measure drug or antibody level in participat treated with the specified TNF-α inhibitor and that at least one source for the specified combination of the test kit and TNF-α inhibitor has been identified in the cost-effectiveness systematic reviews.

ADL: adalimumab; ETN: etanercept; IFX: infliximab

Notes:

* The type of Sanquin test kits used in these studies (MabTrack or those used by Sanquin Diagnostic Services) was not reported.

[1] Gavan (2017) (personal communication), Laine and colleagues (2016)[2] and Ucar 2017

[2] Krieckaert and colleagues (2015)[1] and Laine and colleagues (2016)[2]

[3] Laine and colleagues (2016)[2] and Ucar and colleagues (2017)

[4] Indicates that a test for total anti-drug antibodies is also available (total anti-drug antibodies include both unbound, i.e. free, antibodies and those bound to TNF-α inhibitor)

[5] Gavan (2017) (personal communication) and Laine and colleagues (2016)[2]

[6] Laine and colleagues (2016)[2]

Both Krieckaert and colleagues and Laine and colleagues reported that TDM was cost saving compared to standard care, based on follow up periods of up to three years. Krieckaert and colleagues reported a formal cost per QALY analysis in which TDM dominated standard care in the base-case scenario in 72% of simulations. The ICERs are arguably somewhat meaningless given the small QALY differentials involved. In a range of sensitivity analyses, a net loss of QALYs with respect to the intervention was associated with drug level cut offs, the use of EULAR good response as an outcome, or the use of biologicals other than TNF inhibitors. With regard to UK clinical practice, Krieckaert and colleagues modelled testing at the 28th week, and considered dose reduction by prolongation of the interval between drug administrations in responders with high levels of ADL (DL >12mg/l). The Assessment Group (AG) is aware, however, that in the UK, there are variations as to when treatment decisions in people with RA on biologics are made. In responders to anti-TNF inhibitors, decisions could be made either nine to 12 months after

treatment initiation, or adjusted approximately two years after the initiation of biologic therapy. In non-responders, however, testing may be considered earlier to detect whether non-response to biologics is due to low drug levels or the presence of anti-drug antibodies.

Laine and colleagues (2016) did not report a cost per QALY analysis, although the authors attempted to analyse the frequency and cost impact of non-testing with regard to inappropriate treatment decisions (e.g. continuation of ineffective therapy). The assumption was made that participants in the routine practice arm would typically experience three months' delay in receiving optimal treatment compared to participants in the intervention arm. This was justified based on the typical follow up intervals of participants in Finland. Of note, is that in Finland, anti-drug antibody levels of at least 30 U/mL rather than 12 U/mL are considered clinically significant.

Both the INGEBIO study (Arango and colleagues, 2017, n = 169) and Pascual-Salcedo and colleagues (2013) (n = 88) recruited mixed populations featuring respectively 37% and 50% of participants with a diagnosis of RA. In addition, there were only limited details of the input parameters and analysis, specifically:

- No details of utility values or incremental QALY outcomes were provided.
- The studies did not consider specific test-based treatment algorithms.
- Pascual-Salcedo and colleagues (2013) did not specify which ELISA test kits were used in their study.

Furthermore, the allocation of participants to groups in the INGEBIO was site dependent, and physicians were not obliged to follow any particular algorithm with regard to treatment. However, the statistical analysis plan was not documented, and the assumption of independence of observations may not be appropriate. Therefore the statistical significance of the reported results may be insecure.

The recent study by Gavan (2017) perhaps most closely matches the decision problem. In this study, modelling was based on patient data from the BSRBR-RA register which is the main source of evidence on the use of biologics in people with RA in the UK. Furthermore, the research questions, addressed in Gavan 2017, are most relevant to the decision problem considered in this report. Gavan however did not consider any specific test kit, and only ADL treatment was modelled as firstline.

Of Gavan's three stated research questions (refer to Gavan p. 59), namely:

- Research Question 1: *What was the existing economic evidence for stratified medicine in RA?*
- Research Question 2: *How were treatment decisions with biologic therapies made for patients with RA in current practice in England?*
- Research Question 3: *Are treatment decisions stratified by ADAb and drug level testing, for patients with RA in England, a relatively cost-effective use of healthcare resources?*

Questions 1 and 2 have to some extent also been addressed by the searches and consultations for this review. However, Gavan (2017) considered any test-based strategy of a biomarker to stratify treatment decision with any pharmacological therapy whereas the current review focuses on ELISA testing used to monitor response to TNF-α inhibitor treatment. Question 3 aligns closely with the decision problem for this appraisal. Although Gavan (2017) pointed out that there was a high degree of decision uncertainty and also reported an expected value of perfect information (EVPI) estimate of £7,000,000. The decision uncertainty was based around the cost of testing and test accuracy.

Based on these searches, it is clear that further exploration of this question, including de-novo modelling, would be appropriate. Sufficient prior evidence with regard to the entirety of the decision problem and/or UK populations upon which to base decision making, clearly does not exist, especially given discrepancies in the conclusions of the studies presented. Of these studies, only Gavan (2017) could be considered to be of sufficient quality. However, we found no evidence with regard to either the use of test kits for certolizumab pegol or golimumab treatments, and no studies were identified evaluating IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits or MabTrack ELISA kits.

3.6 Conclusions

The results of the cost-effectiveness systematic review conducted in this study indicate limited evidence on the cost-effectiveness of therapeutic drug monitoring in people with RA. Despite a comprehensive search of the literature, only five studies have been identified. Two (out of five) TNF testing kits from the NICE scope (Promonitor and Sanquin) and three (out of five) TNF inhibitors (ADL, ETN, and IFX) have been assessed in the selected studies.

Two out of five identified studies were reported in abstract format only and therefore limited detail was reported.

4 Independent economic assessment

4.1 Methods

4.1.1 Summary of available evidence

A systematic review of economic evaluations was completed (refer to Section 3). Table 38 provides an overview of those tumour necrosis factor-alpha (TNF-α) treatments and enzyme linked immunosorbent assay (ELISA) kits from the NICE scope which were considered in the studies identified in the cost-effectiveness systematic review.

A systematic review of clinical effectiveness evidence was conducted (refer to Section 2). Table 39 summarises which combinations of treatments and ELISA kits were used in the included studies.

Table 39: Clinical-effectiveness evidence relevant to specific combinations of TNF-α inhibitors and test kits from the NICE scope

		Promonitor	IDKmonitor	LISA-TRACKER	RIDASCREEN	Sanquin*
ADL	*drug*	✓[1]	X	X	X	✓[2]
	antibody	✓[3]	X[4]	X	X	✓[2]
ETN	*drug*	✓[5]	X	X		✓[2]
	antibody	✓[6]	X	X		✓[2]
IFX	*drug*	✓[7]	X[4]	X	X	✓[8]
	antibody	X	X	X	X	✓[8]
GLM	*drug*	X	X	X		X
	antibody	X	X	X		
CTZ	*drug*			X		X
	antibody			X		

Key:

X Indicates availability of a test to measure drug or antibody level in people treated with the specified TNF inhibitor and that no studies have been identified in the clinical-effectiveness systematic review, reporting on using TDM for the specified test kit and TNF inhibitor

✓ Indicates availability of a test to measure drug or antibody level in people treated with the specified TNFi and that at least one source for the specified combination of the test kit and TNF inhibitor has been identified in the clinical-effectiveness systematic review.

ADL: adalimumab; CTZ: certlizumab pegol; ETN: etanercept; GLM: golimumab; IFX: infliximab; TDM: therapeutic drug monitoring; TNF-α: tumour necrosis factor-alpha

Notes:

* The type of Sanquin test kits used in these studies (MabTrack or those used by Sanquin Diagnostic Services) was not reported.

[1] Arango and colleagues 2017, Ucar and colleagues 2017 and Gorostiza and colleagues 2016; Chen and colleagues 2016; Inciarte-Mundo and colleagues 2016, Rosas and colleagues 2013 and Senabre Gallego and colleagues 2017

[2] Paredes and colleagues 2015, Paredes and colleagues 2016 and Pascual-Salcedo and colleagues 2013

[3] Arango and colleagues 2017, Ucar and colleagues 2017 and Gorostiza and colleagues 2016; Chen and colleagues 2016, Rosas and colleagues 2013

[4] Indicates that a test for total anti-drug antibodies is also available (total anti-drug antibodies include both unbound, i.e. free, antibodies and those bound to TNF-α inhibitor)

[5] Inciarte-Mundo and colleagues 2016, Rosas and colleagues 2013, and Senabre Gallego and colleagues 2017

[6] Rosas and colleagues 2013

[7] Inciarte-Mundo and colleagues 2016

[8] Lopez-Casla and colleagues 2013, Paredes **and** colleagues 2015 and Paredes and colleagues 2016

As shown in Table 39, no clinical-effectiveness evidence related to IDKmonitor, LISA-TRACKER, RIDASCREEN and MabTrack tests has been identified. In studies which used Sanquin test kits, the type of kits was not reported. For two drugs from the NICE scope, golimumab (GLM) and certolizumab pegol (CTZ), no studies were found that investigated the use of therapeutic drug monitoring (TDM) in people with RA treated with TNF-α inhibitors. In those studies where antibody testing was conducted, the type of testing (*reflex* or *concurrent* testing) was not reported. In the clinical-effectiveness systematic review, no studies reporting on the use of ELISA testing in people with RA receiving biosimilar products were identified.

All of the included studies except that reported by Lopez-Casla and colleagues (2013)[62] included people in remission or with low disease activity (LDA), while Lopez-Casla and colleagues (2013)[62] considered a mixed population of people with RA comprised of primary and secondary non-responders who had developed clinical inefficacy to infliximab (IFX) (refer to Section 2.3.3 for further details).

In two out of eight studies the study populations were mixed, with 37% of people with RA in the INGEBIO study and 49% in Pascual-Salcedo and colleagues (2013).[59] Moreover, populations considered in the selected studies were relatively small, with the only exception being the INGEBIO study which had a (*mixed disease*) population including 169 participants (Section 2.3.2.1).

In the INGEBIO STUDY which compared test vs. no-test treatment strategies (Section 2.3.3.1), physicians were not obliged to follow any therapeutic algorithm based on TDM results but could use testing to alter doses based on their clinical judgement in participants from the intervention arm. The study was conducted in Spain. The longest follow-up of 18 months was reported by Arango and colleagues (2017).[15] Some of the clinical outcomes are shown in Table 40.

Table 40: Clinical outcomes and follow-up period from Ucar and colleagues (2017) and Arango and colleagues (2017)

Outcome	Ucar and colleagues 2017 Intervention arm	Ucar and colleagues 2017 Control arm	Arango and colleagues 2017 Intervention arm	Arango and colleagues 2017 Control arm
Proportion of patients with tapered dose, %	35.8%	36.7%	35.7%	34.6%

Outcome	Ucar and colleagues 2017 Intervention arm	Ucar and colleagues 2017 Control arm	Arango and colleagues 2017 Intervention arm	Arango and colleagues 2017 Control arm
Rate of flares per patient-year	0.463[1]	0.639[1]	0.463[1]	0.639[1]
Mean duration of remission/LDA, days	344	329	460.2	475.2
Mean follow-up, days	499	505	530.8	544.6

Key: LDA: low disease activity

Note:

[1]The rate of flares per patient-year reported in Ucar and colleagues 2017 is the same as in Arango and colleagues 2017 (even though these sources reported outcomes for different follow-up periods)

Source: Ucar and colleagues (2017)[11] and Arango and colleagues (2017)[15]

The authors reported the mean cost of adalimumab (ADL) treatment per patient-year, and mean quality-adjusted life years (QALYs) (based on EQ-5D-5L) accrued over the observed period in the intervention and control arms. Since the study was reported in the abstract form only, it was not clear how the mean QALYs were estimated.

4.1.1.1 Search for additional effectiveness evidence

Due to the lack of randomised controlled trial (RCT) evidence on the effectiveness of the tests that are defined within the NICE scope, a systematic review of the literature was conducted to identify RCTs evaluating any tests used to monitor anti-TNF-α treatment in people with RA. The aim of this search was to identify any evidence on the effectiveness of any strategies of treatment monitoring that could be used to inform scenario analyses for the modelling.

Searches were carried out in MEDLINE, MEDLINE In-Process, Embase, The Cochrane Library and Web of Science. Searches were limited to RCTs and carried out in October 2018. The search strategy is provided in A3.2.

A total of 1,418 hits were identified and independently screened by two reviewers versus the inclusion criteria shown in Table 41. No relevant papers were identified.

Table 41: Inclusion criteria

Criteria	Specification
Population	As for the clinical-effectiveness systematic review (see Section 2.2.1.1)
Interventions	Any test outside of the scope for monitoring patients receiving TNF-α inhibitors (ADL, ETN, IFX, CTZ, GLM).
Comparator	Current practice (i.e. no testing)
Outcomes	As for the clinical-effectiveness systematic review (see Section 2.2.1.4)

Criteria	Specification
Study design	RCT

Key: ADL: adalimumab; CTZ: certolizumab pegol; ETN: etanercept; GLM: golimumab; IFX: infliximab; RCT: randomised controlled trial

4.1.2 Choice of modelling approach

The choice of the modelling approach in this assessment was primarily driven by the availability and quality of the evidence identified in the clinical-effectiveness systematic review (refer to Section 2). Other factors influencing the choice of approach included: the multifactorial nature of decisions to adjust treatments in people with RA,[50] and the recent changes in the biologics market, which contributed to the uncertainty in the prices of the various biologics and their uptake within the UK (Section 4.1.9.1.3).

A simplified modelling approach, a threshold analysis, was therefore chosen to address the decision problem. Although not a formal cost-effectiveness study, this approach allowed the estimation of the cost of TNF testing, at which the test-based treatment has zero net monetary benefit (NMB), while taking into consideration the major components of differential costs and QALYs.

NMB represents the value of an intervention in monetary terms when a willingness-to-pay (WTP) threshold for a unit of benefit (e.g. QALY) is known. NMB is estimated by first assuming a WTP threshold (e.g. £20,000 or £30,000 per QALY) and then calculating thse NMB as follows:

$$incremental\ benefit\ x\ threshold - incremental\ cost,$$

where $incremental\ cost$ and $incremental\ benefit$ represent differential costs and QALYs for the health technologies under consideration.

The differential costs of drug acquisition, drug administration, and disease management were considered; the latter comprised the costs of managing flares and adverse events (AEs), and the costs associated with managing different health states. QALYs were estimated from the rates of flares and AES, and the duration of remission in the intervention and control arms. Refer to Section 4.1.4 for further details on the modelling approach.

In addition to the threshold analysis, incremental cost-effectiveness ratios (ICERs) were estimated using the list prices of the originator products and their biosimilars (Table 50) assuming similar clinical effectiveness across the TNF-inhibitors and similar performance of the Promonitor test kits used for measuring the drug and antibody levels of the TNF

inhibitors. Estimates of the cost of testing were based on Jani and colleagues (2016)[53] and clinical advice. Clinical outcomes from the INGEBIO study were utilised in all of the analyses.

4.1.3 Analyses

4.1.3.1 Patients in remission/low disease activity

The clinical and economic effect of ADL tapering in people with RA in remission/LDA was evaluated in the INGEBIO study (Arango and colleagues 2017, Ucar and colleagues 2017 and Gorostiza and colleagues 2016),[11,15,64] with the longest follow-up reported in Arango and colleagues 2017 (530.8 and 544.6 days, on average, in the intervention and control arms, respectively). In this study, ELISA testing of drug and anti-drug antibody levels was compared against usual practice of clinically driven monitoring alone. Arango and colleagues (2017) reported improved HRQoL (EQ-5D-5L) and lower average cost of Humira® per patient-year in the intervention group. Of note, the between-arm difference of -0.176 in the rate of flares per patient-year was *not statistically significant* (95% confidence interval [CI]: -0.379 to 0.0289) (Arango and colleagues 2017);[15] flare was defined as an increase in DAS28 >1.2, or an increase in DAS28 >0.6 if the current DAS28 ≥3.2.[67]

The INGEBIO study included a mixed population of 169 people treated with ADL for RA, psoriatic arthritis (PsA) and ankylosing spondylitis (AS). A total of 63 people with RA (30 in the intervention group and 33 in the control group), which constituted 37% of the total study population.[11] ADL frequency was adjusted based on physician criteria (i.e. physicians were not required to follow any therapeutic algorithm based on TDM results but could use tests to alter doses based on their clinical judgement). Drug and antibody levels were measured using Promonitor test kits (Table 12). See Section 2.3.3.1 for a detailed description of the study.

The clinical outcomes reported in Ucar and colleagues (2017)[11] (Table 40) were incorporated in this economic analysis in order to evaluate the cost of drug and antibody testing, at which the addition of ELISA testing to usual clinical practice would result in zero NMB. Importantly, Ucar and colleagues (2017)[11] reported clinical outcomes for a slightly shorter follow-up period compared with Arango and colleagues (2017)[15] (Table 40). However, these outcomes were based on the intention-to-treat (ITT) population while the estimates reported in Arango and colleagues (2017)[15] did not include 19 participants lost to follow-up. It may be argued that, because *participants lost to follow-up* often have a different prognosis than those who complete the study, excluding such participants may bias the results. Importantly, estimates from Ucar and colleagues (2017)[11] were directly applied in the

analysis as quantitative synthesis of evidence related to people in remission/LDA identified in the clinical-effectiveness systematic review, was not possible (Section 2.2.5).

Since the patent for the ADL originator product, Humira®, expired in October 2018, and the acquisition costs for the ADL biosimilars were not known to the AG at the time of writing (Table 50), the annual acquisition cost was varied from £1,000 to £9,180 per patient-year in the threshold analysis. The latter represents the annual cost of Humira® assuming the dose of 40 mg every two weeks delivered by subcutaneous injection using a pre-filled pen and the NHS indicative price from the British National Formulary (BNF) (Table 50).

The other major assumptions were as follows, with further details in ADL dose tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).

- Dose is tapered in a proportion of people in each arm at the start of simulation.
- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored *in all people* on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), and, therefore, the administration cost is zero.
- The costs associated with flare management are:
    - £423 per flare for diagnostic investigations
    - £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.
- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.
- The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
- The duration of AE is 28 days.
- The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 42:

- ADL dose tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).
- Dose is tapered in a proportion of people in each arm at the start of simulation.
- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored *in all people* on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), and, therefore, the administration cost is zero.
- The costs associated with flare management are:
    - £423 per flare for diagnostic investigations
    - £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.
- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.
- The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
- The duration of AE is 28 days.
- The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 42: Model assumptions in the analyses with people in remission/low disease activity

Assumption	Estimate	Source	Relevant table/ sections in the report
Dose tapering strategy	Spacing: from 40 mg of ADL every 2 weeks to 40 mg every 3 weeks	1st dose reduction in the Exeter biologic clinic recommendations (Appendix 5)	Section 4.1.9.1.5
Proportion of patients on tapered dose:			
Intervention	35.8%		Table 40
Control	36.7%		Table 40

Assumption	Estimate	Source	Relevant table/ sections in the report
Proportion of flared patients in whom the full dose is restored	100%	Exeter biologic clinic recommendations	Appendix 5
Mean duration of remission (days)			
Intervention	344		Table 40
Control	329		Table 40
Mean follow-up (days)	505	As in the control arm (Ucar and colleagues, 2017)	Table 40
Acquisition costs (per patient-year): Humira®			
Full dose[1]	£9,187	BNF	Section 4.1.9.1.3
Tapered dose	£6,125	BNF, Exeter biologic clinic recommendations	Appendix 5
Flared patients[2]	£9,187	BNF, Exeter biologic clinic recommendations	Appendix 5
Treatment wastage on full dose (per patient-year)	£370	Clinical advice	Section 4.1.9.1.6
Administration cost for Humira® (ADL) (per patient-year)[2]	£0	Clinical advice	Section 4.1.9.1.7
Cost of flare management [3, 4]	£423/per flare	Cost of diagnostic investigations (Maravic and colleagues, 2005[4])	Section 4.1.9.1.19
	£68/month	Monthly cost of treatment (excluding DMARDs) (Maravic and colleagues, 2005[4])	Section 4.1.9.1.19
Cost of managing health states (per patient-year) [5]			
Remission	£11,409	Barbieri and colleagues (2005),[5] Radner and colleagues (2014),[6]	Section 4.1.9.1.16
LDA/active disease	£18,889	National Schedule of Reference Costs 2017-18[7]	Section 4.1.9.1.16
Cost of managing AEs (per infection)	£1,622[6]	TA375[8]	Section 4.1.9.1.20
Utilities			
Remission	0.717	Estimated from HAQ scores for different HAQ	Section 4.1.9.2.1

Assumption	Estimate	Source	Relevant table/ sections in the report
LDA/active disease	0.586[7]	bands reported by Radner and colleagues (2014)[6]	Section 4.1.9.2.1
Disutility of flare	0.140	Markusse and colleagues, 2015[9]	Section 4.1.9.2.2
Disutility of AEs	0.156	TA375,[8] Oppong and colleagues (2013)[10]	Section 4.1.9.2.3
Flare rate			
Intervention	0.463	Ucar and colleagues 2017[11]	Section 4.1.8.1.1
Control	0.639	Ucar and colleagues 2017[11]	Section 4.1.8.1.1
Mean time to first flare (days)			
Intervention	208.07	Derived from Kaplan-Meier estimates (from the INGEBIO study) of time to first flare, provided by Ucar and colleagues (personal communication, 9 September, 2018)	Section 4.1.8.1.3
Control	189.32		Section 4.1.8.1.3
Flare duration (days)[8]	7	TA375[8]	Section 4.1.8.1.2
Rate of AEs			
Patients on full ADL dose	3/100 patient-years	Senabre Gallego and colleagues (2017)[12]	Section 4.1.8.2.1
Patients on reduced ADL dose	2/100 patient-years[9]	Singh and colleagues (2015)[13]	Section 4.1.8.2.1

Assumption	Estimate	Source	Relevant table/ sections in the report
Duration of AE (days)	28	TA375,[8] Oppong and colleagues (2013)[10]	Section 4.1.8.2.2

Key: ADL: adalimumab; AE: adverse event; BNF: British national Formulary; HAQ: health assessment questionnaire; HAD: high disease activity; MDA: moderate disease activity; OR: odds ratio; PPP: purchasing power parities; RA: rheumatoid arthritis; RCTs: randomized controlled trials; TA: technology appraisal

Notes:

[1] Assuming 40 mg every two weeks by subcutaneous injection using pre-filled pen, and NHS indicative price from the BNF.

[2] The mean time to first flare was estimated from additional evidence (Kaplan-Meier curves for time to first flare) from the INGEBIO study provided by Ucar and colleagues (2007)[11] (poster, personal communication).

[3] The estimates were derived from the costs of managing a flare in a hypothetical person with a 10-year history of RA in the French setting. The costs were converted to pound sterling based on PPP and inflated to 2017-18 prices using the healthcare price index (Section 4.1.9.1.1).

[4] The estimates from Maravic and colleagues (2005)[4] do not include the cost of rheumatology appointments.

[5] The costs of managing health states were included by HAQ-dependency, i.e. by assigning an annual cost to mutually exclusive HAQ intervals.

[6] The estimate of £1,479 per patient-year from the source was inflated to 2017-18 prices using the healthcare price index (Section 4.1.9.1.11).

[7] The estimate was computed from a weighted average HAQ score for the LDA, MDA, HDA health states reported by Radner and colleagues (2014)[6] and mapped to EQ-5D values following Malottki and colleagues (2011)[14] (Section 4.1.9.2.1).

[8] This estimate was used for calculation of QALYs only since it was assumed that the ADL dose in people with flares is switched back to the full dose indefinitely.

[9] Based on OR=1.31 for standard-dose biologics in people with RA reported by Singh and colleagues (2015).[13] The OR estimate was obtained in a Bayesian network meta-analysis (using a binomial likelihood model) of 11 published RCTs (n=4,788) to assess the risk of serious infections in anti-TNF-biologic-experienced people with RA.

Of note, in the primary analysis, QALYs were estimated based on heath-state utilities as well as disutilities of flares and AEs. It was assumed (based on Smolen and colleagues, 2017,[76]) that people in any health state (i.e. in remission, LDA and active disease) can experience flares (Section 4.1.8.1). Utilities for the mixed disease population in the INGEBIO study were assumed to be the same as those for the population of people with RA only since no evidence on HRQoL directly relevant to the population considered in INGEBIO has been identified.

When modelling the effect of AEs on HRQoL and costs, the Assessment Group (AG) adopted the approach used in TA375 - it was assumed that only serious adverse events (serious infections in particular) would carry a significant cost and disutility burden (p. 381, TA375[8]). This was supported by the opinion of our clinical experts.

Mortality associated with RA was not modelled and no discounting was applied to the costs and outcomes due to the short-term time horizon of about 18 months.

4.1.3.2 Primary and secondary non-responders

The only study identified in the clinical-effectiveness systematic review which considered non-responders to TNF-α inhibitors (Lopez-Casla and colleagues 2013)[62] was a prospective cohort study conducted in Spain to assess whether increasing infliximab (IFX) dose was an

effective therapeutic option (refer to Section 2.3). The study included 36 people with RA (23 primary and 13 secondary non-responders) who had developed clinical inefficacy to IFX and in whom IFX dose increase was implemented. ELISA test kits by Shandon Diagnostic Services were used to measure trough and antibody levels.

This study did not compare a test-based treatment strategy with standard clinical practice. It was also considered low quality as it was a *retrospective observational study* judged to be *at a* moderate *risk of bias* due to missing data or confounding factors, *with a relatively small number of patients* (refer to Section 2.3.4). As a result, data reported by Lopez-Casla and colleagues (2013)[62] did not serve as the basis for the economic analysis considering non-responders.

Due to the lack of relevant data identified in the systematic review of clinical effectiveness, the cost-effectiveness of TNF-testing in non-responders could not be evaluated.

4.1.4 Model structure

The cost of TNF testing, under which the treatment strategy based on test results and clinical judgement has zero NMB, was estimated in the following way:

$$Total\ cost\ of\ testing = ICER\ threshold * \Delta QALYs - \Delta Costs,$$

where $Total\ cost\ of\ testing$ was comprised of the cost of the expected resource use and costs associated with testing patient samples to monitor drug trough and anti-drug antibody levels (refer to Section 4.1.9.1.8); and, $ICER\ threshold$ represents the *NICE* cost-effectiveness threshold of either £20,000 per QALY or £30,000 per QALY gained. Threshold analyses were conducted for both thresholds.

The costs incurred in each arm were estimated as follows:

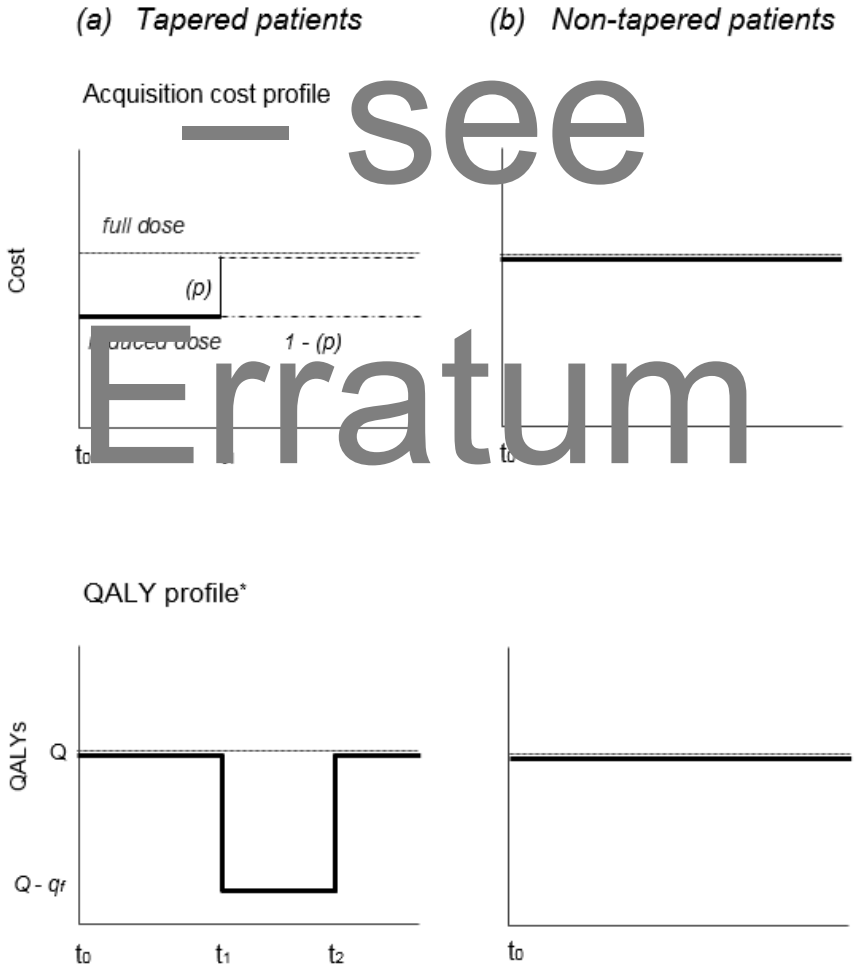
$$Costs = acquisition\ cost + administration\ costs + cost\ of\ managing\ health\ states + cost\ of\ managing\ flares + costs\ of\ managing\ adverse\ events$$

QALYs were derived as follows:

$$QALYs = duration\ of\ remission * utility\ score\ for\ remission + duration\ of\ active\ disease * utility\ score\ for\ active\ disease - average\ duration\ of\ flare * rate\ of\ flare * disutility\ of\ flare - average\ duration\ of\ adverse\ event * rate\ of\ adverse\ event * disutility\ of\ adverse\ event$$

Based on Gavan (2017),[35] Figure 8 shows a graphical illustration of the cost and QALY profile depending on whether the dose is tapered or not. The figure shows changes in the acquisition cost and QALYs *due to flares* over time. Note that, for the sake of clarity, the other components of the total costs and QALYs considered in the analysis, are not depicted here.

Figure 8: Acquisition costs and QALY change in tapered and non-tapered patients due to flare



Notes:
($t_1 - t_{0)}$ is the time on tapered dose; ($t_2 - t_{1)}$ is the duration of flare; q_f is the disutility of flare.
* Change in QALYs due to flare
Source: Based on Gavan (2017)[35]

In this scenario, all patients had their drug levels tested at (t_0), and in some patients the dose was tapered (Figure 8 (a)). A proportion (p) of patients on tapered doses were assumed to flare at (t_1), which prompted treatment to revert to its original dose. In those patients who flared the disutility of flare, q_f, was applied for the duration of flare, ($t_2 - t_1$). In non-tapered patients (Figure 8 (b)), the acquisition cost was based on the cost of the full dose of adalimumab. It was assumed that these people do not experience flares.

The impact of flares on costs and QALYs used by Gavan (as shown on Figure 8) was used in the model. It was also assumed (based on Smolen and colleagues, 2017[76]) that flares could occur *in any health state* (i.e. in remission, LDA and active disease) (Section 4.1.8.1).

4.1.5 Population

The population modelled were people in remission/LDA. Table 43 presents baseline characteristics of participants included in the INGEBIO study, used for model parameterisation, along with the characteristics of people with RA, *responding to biologics,* from the British Society for Rheumatology Biologics Register (BSRBR).

Table 43: Patient baseline characteristics

Study	# RA patients	Age	% females	Disease duration, years	Treatment history	Concomitant treatments	Disease state
INGEBIO	Mixed pop: 63 people with RA of total 169 participants	53.6[1]	42%[1]	Median=10	NR	MTX[2] – 76.7%	77% people in remission, 23% LDA (at baseline)
BSRBR data for responders[3]	10,186	56	76.3%	Mean=13 (years at the time of initiation of 1st biologic)	Mean=3.90 (previous DMARDs)	*NR*[4]	30.6% – good responders

Key: BSRBR: British Society for Rheumatology Biologics Register; DMARDs: disease modifying anti-rheumatic drugs; LDA: low disease activity; MTX: methotrexate; NR: not reported; RA: rheumatoid arthritis; TNF: tumour necrosis factor

Notes:

[1] Weighted average across treatment arms

[2] Patients concurrently receiving anti-inflammatories or MTX are more likely to respond to anti-TNF (Dennison and colleagues, 2016)[77]

[3] Table 189 (TA375,[8] p.367). Of note, as stated in the source, the BSRBR database contained a very small number of MTX-naïve patients at the time the analysis was performed.

[4] As stated in TA375,[8] (p. 354) the BSRBR database contains a very small number of MTX-naïve patients.

Subgroups

People with RA can be grouped according to three clinical scenarios: primary non-response, secondary non-response and remission. However, with regards to particular characteristics, there are no subgroups for which the tests are expected to significantly vary. Therefore, no subgroups were considered in this assessment.

4.1.6 Interventions and comparators

Due to the paucity of data, not all test kits specified in the NICE scope could be evaluated in this economic analysis. In particular, no economic analysis relevant to IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits was conducted. The only test kits considered were Promonitor kits for measuring ADL trough and antibody levels (Table 39).

The comparator was standard care where treatment decisions were based on clinical judgements and other measures (such as DAS28), i.e. without the use of TDM.

4.1.7 Perspective, time horizon and discounting

The costs and resource use in this analysis were considered from the perspective of the NHS and Personal Social Services.[78]

The time horizon was 18 months as defined by the observational period in the INGEBIO study.

Cost and health outcomes were not extrapolated into the future as, due to the lack of long-term evidence, external validation of extrapolated outcomes was not feasible. Therefore, no discounting was applied to estimated costs and QALYs.

4.1.8 Considerations in the development of the independent economic assessment

4.1.8.1 Flares

The concept of flare remains challenging to understand as there are no generally recognised definitions of or well-validated measures for flare in RA (Bykerk and colleagues, 2014[16]). Nevertheless patients, clinicians and scientists commonly resort to this term referring to episodes of worsening disease activity which includes a range of symptoms of different duration and magnitude (Bingham and colleagues, 2009[3]).

Three studies included in the clinical-effectiveness systematic review reported a definition of flare: all used DAS28 >3.2 (refer to Section 2.3.2.1). In the INGEBIO study, a disease flare was defined as an increase in DAS28 >1.2, or an increase in DAS28 >0.6 if the current DAS28 ≥3.2. This followed criteria validated by van der Maas and colleagues (2013)[67] (this information was provided by the authors on request by the AG).

The AG is aware of several RA flare criteria, which have been used in clinical research. For instance, Van der Maas and colleagues, 2013[67] identified six previously published DAS28-based flare criteria and Markusse and colleagues (2015)[9] reported three criteria (Table 44).

Table 44: The definition of flares from Maas and colleagues (2013) and Markusse and colleagues (2015)

Type	DAS28		
	current	previous	increase
Van der Maas and colleagues (2013)[67]			
1	any	NA	> 1.2
	> 5.1	NA	> 0.6
2	any	NA	> 1.2
	≥ 3.2	NA	> 0.6
3	any	NA	> 0.6
	> 3.2	NA	any
4	any	NA	> 1.2
5	> 3.2	NA	any
6	> 2.6	NA	any
Markusse and colleagues (2015)[9]			
A	> 2.4	any	≥ 0.6
Minor B	> 2.4	≤ 2.4	< 0.6
*Major B **	> 2.4	≤ 2.4	≥ 0.6

Key: DAS28: disease activity score in 28 joints; NA, not applicable

Smolen and colleagues (2017)[76] compared people with RA treated with ETN, recruited in the PRESERVE trial, who did or did not have flares. In the trial, flare was defined as either *loss of LDA* with/without DAS28 change of 0.6 or *relapse* (DAS28 >5.1 or DAS28 >3.2 at ≥2 time points).

The variation in flare criteria reported in the literature was confirmed by the clinical advisors to be consistent with clinical practice.

4.1.8.1.1 Rate of flares

Annual per patient rates of flares in the intervention and control groups were reported in Ucar and colleagues (2017)[11] and Arango and colleagues (2017),[15] (refer to Table 40 and Section 2.3.3.1). These estimates were used in the primary and scenario analyses, respectively.

4.1.8.1.2 Duration of flare

There is a substantial heterogeneity in the duration of flare. Dr Meghna Jani advised that flare may last from two to three days up to two to three months depending on severity.

The duration of flares was estimated in the dynamic cohort Brigham RA Sequential Study (BRASS), which included 1,105 people with established RA who had received usual care at the Brigham and Women's Hospital in Boston, US (Bykerk and colleagues, 2014[16]) (Table 45).

Table 45: Flare duration in the BRASS study

	Estimates		
Duration, days	<7	7–13	≥14
Proportion of patients, %	57%	14%	30%

Key: BRASS: Brigham RA Sequential Study
Source: Bykerk and colleagues, 2014[16]

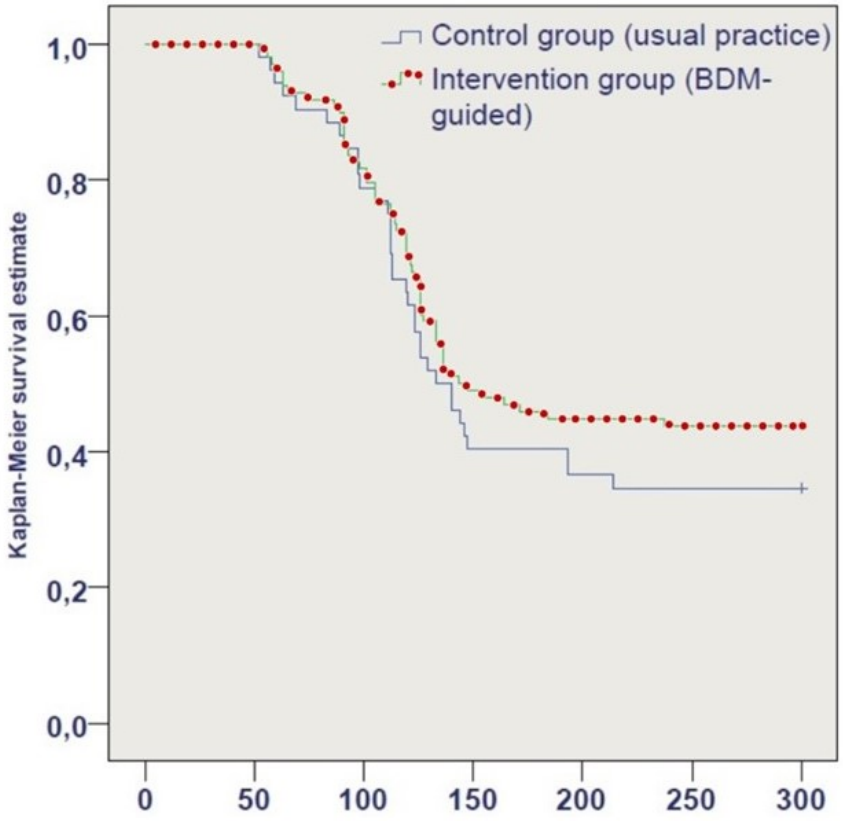
The estimate of seven days was adopted in the primary analysis. This was consistent with the estimate used in TA375.[50] In a scenario analysis, the effect on the results of a longer duration of flare, 19 days, was examined. This estimate was a weighted average of data reported in the BRASS study[16] (Table 45), and the estimates provided by Dr Meghna Jani.

4.1.8.1.3 Time to first flare

Arango and colleagues (2017)[15] and Ucar and colleagues (2017)[11] reported *median time* to first flare observed in the intervention and control arms of the INGEBIO trial. However, according to the NICE *Guide to the Methods of Technology Appraisal,*[78] *mean estimates* should be utilised in economic analyses of health interventions.

The mean estimates for the intervention and control arms were derived from Kaplan-Meier (KM) curves for the time to first flare estimated in the INGEBIO study and reported in an additional source (a poster presentation) provided by Ucar and colleagues(2007)[11] (provided to the AG on request in September, 2018) by using the area under the curve (AUC) approach. The KM estimates were available for 300 days (Figure 9), and were extrapolated for the duration of follow-up reported in Ucar and colleagues (2017)[11] (Table 40) and Arango and colleagues (2017)[15] (Table 40). Since the proportions of participants on tapered dose in the intervention and control groups levelled at around 240 days after dose tapering, it was assumed that these proportions remained the same until the end of the observational periods in Ucar and colleagues (2017)[11] and Arango and colleagues (2017)[15] (i.e. no parametric model fitting was performed). Estimates of the mean time to first flare are shown in Table 46.

Figure 9: Kaplan-Meier estimates from INGEBIO study



Key: BDM: biologic drug monitoring

Table 46: Mean time to first flare (days)

	Intervention	Control
Primary analysis (Ucar and colleagues 2017)[11]	208.07	189.32
Scenario analysis (Arango and colleagues 2017)[15]		

4.1.8.2 Serious adverse events

When modelling the effect of AEs on patients' HRQoL and costs, the AG adopted the approach used in TA375 – it was assumed that only serious adverse events (serious infections in particular) would carry a significant cost and disutility burden (p. 381, TA375[50]). This assumption was considered appropriate in the opinion of the AG's clinical advisors.

4.1.8.2.1 Rate of serious adverse events

One study includedin the clinical-effectiveness systematic review, Senabre Gallego and colleagues (2017),[12] reporting AEs of anti-TNF therapies such as infections. This study recruited 39 people with RA who had achieved remission. The finding showed that one

participant (3%) had septic arthritis (serious infectious arthritis) associated with anti-TNF therapies (ADL or ETN) in the one-year follow-up period (Table 47).

Since the identified evidence was limited, additional searches were conducted to identify published estimates of the rate of serious infection in people with RA treated with biologics. Data reported by Lahiri and colleagues (2015),[79] indicated a time-dependent increase in the risk of serious infections, with the maximum risk within the first six months of therapy and a gradual decline thereafter. This time-dependent decrease in risk can be attributed both to 'depletion of susceptibles' (i.e. high-risk participants dropping out of the anti-TNF cohort due to death, stopping therapy or loss to follow-up), accounting for two-thirds of the observed difference, and to improvement in the inherent infection risk by an improvement in the functional status and a decrease in the dose of glucocorticoid (GC). Consultation with clinical advisors confirmed that serious infections in people with RA from the population of interest are relatively rare.

According to Bruce at al. (2016),[80] the risk of pneumocystis jirovecii *pneumonia* (PJP) in people with RA from the BSRBR register, treated with TNF-α inhibitors, was low, with an incident rate of 2.0/10 000 person-years follow-up (95% CI 1.2 to 3.3) (Table 47).

The rate of tuberculosis in people with RA from the BSRBR register treated with anti-TNF therapy, estimated by Dixon and colleagues (2010),[81] was higher for ADL (144 events/100 000 person-years) and IFX (136/100 000 person-years) when compared to ETN (39/100 000 person-years) (Table 47).

The rate of serious adverse events reported in Burmester and colleagues (2017)[82] was 4.7 per 100 patient-years (Table 47). This estimate was derived from 15,132 people with RA exposed to ADL in 28 global clinical trials. A serious adverse event was defined as fatal or immediately life-threatening; requiring hospitalisation or prolonged hospitalisation; resulting in persistent or significant disability/incapacity, congenital anomaly or requiring medical or surgical intervention to prevent a serious outcome. Baseline characteristics of participants considered in this study were as follows: mean age of 53.5 years, 78.8% female, a mean disease duration of 9.1 years, and 16.5% and 10.9% of people on treatment for >2 years and >5 years, respectively.

The rate of serious infections adopted in TA375,[50] 35 per 1,000 patients, was based on Singh and colleagues (2011).[83] It was assumed that the rate of serious infections was independent of the biological DMARDs (bDMARDs) used (i.e. all biologic therapies were assumed to have similar safety profiles). A sensitivity analysis was conducted setting the risk

of AEs for ETN, ADL, and IFX to 0.03767, 0.04075 and 0.04075, respectively, based on the Galloway BSRBR data (data were not available for other biologics from this BSRBR analysis).

Table 47: Serious adverse events in RA patients treated with anti-TNF inhibitors

Source	Type of serious infections	Population	Estimate
Singh and colleagues (2015)[13] (a systematic review)	Serious infection mostly included infections associated with death, hospitalisation, or the use of intravenous antibiotics.	4,788 anti-TNF-experienced people with RA recruited to 11 RCTs during 2005-2013, with mean RA duration of 10.8 years	19/1000
TA375 (based on a systematic review by Singh and colleagues, 2011[83])	Serious infections included opportunistic infections as well as bacterial infections in most studies.	Adults (aged 16 years or older) with any disease (except HIV/ AIDS) included in studies of any of the nine biologics (abatacept (Orencia®), adalimumab (Humira®), anakinra (Kineret®), certolizumab pegol (Cimzia®), etanercept (Enbrel®), golimumab (Simponi®), infliximab (Remicade®), rituximab (Rituxan or Mabthera®) and tocilizumab (Actmera®))	35/1000
Senabre Gallego 2017[12]	Septic arthritis	39 people in clinical remission	1 patients (out of 39) discontinued treatment due to the AE (study FU – 12 months)
Dixon and colleagues (2010)[81]	Tuberculosis	People with RA from the BSRBR register, treated with ADL, ETN or IFX	ADL - 144/100 000 pyrs, ETN - 39/100 000 pyrs, IFX - 136/100 000 pyrs
Bruce at al. (2016)[80]	Pneumocystis jirovecii pneumonia	People with RA from the BSRBR register, treated with anti-TNF ihibitors	2.0/10 000 pyrs (95% CI 1.2 to 3.3)
Burmester and colleagues (2017)[82]	Serious adverse event (defined as fatal or immediately life-threatening; required hospitalisation or prolonged hospitalisation; resulted in persistent or significant disability/incapacity,	15,132 people with RA exposed to ADL in 28 global clinical trials	4.7 per 100 pyrs

Source	Type of serious infections	Population	Estimate
	congenital anomaly or required medical or surgical intervention to prevent a serious outcome)		
Jani and colleagues (2018)[841]	In the HL group: lower (34%) and upper (16%) respiratory tract infections, urinary tract infections (15%), skin infections including shingles (8%)	People from the BSRBR register (safety data), and the Biologics in RA Genetics & Genomics Syndicate (serological samples)	Low/normal DL: 54 (95% CI 30 to 98)[2] per 1000 pyrs; HL DL: 76 (95% CI 55 to 104)[3]

Key: ADL: adalimumab; BSRBR: British Society for Rheumatology Biologics Register; DL: drug level; ETN: etanercept; HL: high level (of dose); IFX: infliximab; pyrs: person-years; RA: rheumatoid arthritis; TNFi: tumour necrosis factor inhibitor

Notes:

1 TNFi drug levels were measured at 3/6/12 months after biologic initiation and stratified as low/normal or high drug levels as per thresholds defined using concentration-effect curves for each drug. The risk of first and total infections within the first year was analysed. Events occurring on drug or within 90 days of last dose were included.

2 Crude rate in patients with low/normal drug level (n=241)

3 Crude rate in patients with high drug level (n=462)

The AE rate for people receiving a full dose of biologics used in the model was 3/100 patient years as reported in Senabre Gallego and colleagues (2017).[12] The odds ratio (OR) reported by Singh and colleagues (2015)[13] (Table 48) was used to estimate the AE rate in tapered patients, which was 2/100 patient-years. The OR for serious infections in people treated with low-dose biologics compared to people treated with standard dose biologics (Table 48) were obtained from a Bayesian network meta-analyses on the risk of serious infections people with RA; they were 0.71 (95% CrI: 0.5, 1.01) and 0.7 (95% CrI: 0.27,1.68) for consistency and inconsistency models, respectively.[85]

Table 48: Comparison of effect estimates from consistency and inconsistency models from Singh and colleagues (2015)

Comparison	Consistency Model*	Inconsistency Model*
LD biologic +/- traditional DMARD vs. SD biologic +/- traditional DMARD, OR (95% CrI)	0.71 (0.50,1.01)	0.7 (0.27,1.68)

Key: CrI: credible interval; DMARDs: disease modifying anti-rheumatic drugs; LD, low dose; OR, odds ratio; SD, standard dose; vs.: versus

Note:

* Dias and colleagues (2011; updated April 2014)[85] NICE technical support document 4.

Source: Appendix 10b, Singh and colleagues (2015)[13]

4.1.8.2.2 Duration of serious adverse events

In TA375[50] serious infections were assumed to persist for 28 days. This estimate was adopted in the primary analysis.

4.1.9 Model parameters

- Parameter values assumed in the primary analysis for people in remission/LDA are shown in ADL dose tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).
- Dose is tapered in a proportion of people in each arm at the start of simulation.
- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored *in all people* on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), and, therefore, the administration cost is zero.
- The costs associated with flare management are:
    - £423 per flare for diagnostic investigations
    - £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.
- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.

- The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
- The duration of AE is 28 days.
- The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 42. The derivation of the parameter values is detailed in the following sections.

4.1.9.1 Resources and costs

Costs considered in the economic evaluation included the costs of testing, the costs of treatments received by people with RA, and healthcare costs. The costs of testing comprised those of the test kits, staff time to perform test and staff training, the cost of testing service and sample transport. Costs were obtained from the British National Formulary (BNF),[86] NHS Reference Costs,[7] from documents provided by test manufacturers, and published and unpublished sources.

4.1.9.1.1 Conversion to GBP

Where conversion from other currencies to GBP was required, IMF purchasing power parity was used to convert within the year (e.g. from 2010 EUR to 2010 GBP), after which inflation was applied. The Campbell and Cochrane Economic Methods Group (CCEMG) – EPPI-Centre Cost Converter was used for the purchasing power parity (PPP) conversion.[87]

4.1.9.1.2 Inflation to 2017-18 prices

Unit costs were inflated to 2017–18 prices by inflating to 2015–16 prices using the Hospital and Community Health Services Pay and Prices Index (Table 49), and then to 2017–18 prices using the average increase in the index for the previous three years (from 2013–14 to 2015–16), with the average rate of 1.1% per annum.

Table 49: Hospital and Community Health Services combined pay and prices inflation indices

Year	Pay and prices (%)
2008-09	3.9
2009-10	0.6
2010-11	3.0
2011-12	2.1
2012-13	1.7
2013-14	1.1
2014-15	0.9
2015-16	1.3

Source: HCHS pay and price inflation, Pay and Price Series - Department of Health[88]

4.1.9.1.3 Treatment costs

The biological medicines market will increase in complexity over the coming months and years as more originator biological medicines lose patent exclusivity and additional biosimilar medicines come to market.[89]

The patent for ADL (known by brand name Humira®), expired on 16 October, 2018. New medications with similar active properties ("biosimilar" versions) will become available in the NHS in the end of 2018 (Table 10). The following ADL biosimilars have already been approved for use in the UK but have not yet launched (as of 30 November, 2018):

- Amgevita® (Amgen)
- Hulio® (Mylan/Fujifilm Kyowa Kirin)
- Hyrimoz® (Sandoz)
- Imraldi® (Samsung Biogen)

According to Regional Medicines Optimisation Committee Briefing,[90] at least two further biosimilars are expected to become available in the UK during 2019: Cyltezo® (from Boehringer Ingelheim) and the second will be brought to the market by Fresenius Kabi.

The NHS has established a working group to provide oversight of implementing the use of best-value ADL using a commissioning framework launched in September 2017.[91] The framework, authored by the NHS' Medicines, Diagnostics, Personalized Medicine Policy Team, proposes that "*at least 90% of new patients be prescribed the best value biological medicine" within three months of the launch of a biosimilar for a given reference product, and that 80% of existing patients be prescribed the "best value" medicine within 12 months of a biosimilar launch*" (NHS England. Commissioning framework for biological medicines [including biosimilar medicines]: Executive Summary, p4).[91]

As for the current uptake of biosimilars in the UK, according to the Medicines Optimisation Dashboard Data published by the NHS England (September 2018 release),[92] 92% and 85% of people who were prescribed IFX and ETN, respectively, are taking biosimilars. However, there are regional variations in the uptake of the different biosimilars.[92]

In the Royal Devon and Exeter NHS Foundation Trust people with RA prescribed IFX or ETN are usually given biosimilars (see Table 10); however, people prescribed golimumab (GLM) or certolizumab pegol (CTZ) are given the originator products. Biosimilars for ADL have become available only recently (Dr Rich Haigh, personal communication). In the Greater Manchester area, the biosimilar Amgevita® is soon to be used in people prescribed

ADL; for people prescribed IFX biosimilars (Inflectra® and Remsima®) are given; and, a biosimilar (Benepali®) is used in some some people prescribed ETN (Dr Meghna Jani, personal communication).

Although the NICE guidance recommends that people with RA patients receive the anti-TNF treatment with the lowest acquisition and administration costs, in practice other non-cost factors such as patient characteristics, hospital characteristics and changes in regional rheumatology clinical guidelines may influence treatment selection (Gavan, 2017).[35]

4.1.9.1.4 Drug acquisition costs

Annual acquisition costs of the TNF-inhibitors, assumed in the cost-utility analyses, were estimated using list prices (in accordance with NICE guidelines[78]) and assuming adherence to standard dosing regimen for each drug (Table 50).

Table 50: Acquisition costs of biologics

Biologic	Dosing regimen	Cost per dose	Cost per year	Additional cost in Year 1
ADL	40 mg, every 2 weeks. In non-responsive patient dose may be increased to 40 mg, once weekly.			
Humira®*		£352.14	£9,187.08	
Amgevita®		NR		
Cyltezo®		NR		
Imraldi®		NR		
Solymbic®		NR		
Hyrimoz®		NR		
Halimatoz®		NR		
ETN	50 mg, once weekly			
Enbrel®*		£178.75 (25 mg/ 0.5 ml)	£9,326.92	
Benepali/ Brenzys		£164	£8,557.29	
Erelzi		£160.88	£8,394.23	
Lifmior®		NR		
CTZ	Loading dose: 400 mg, at Weeks 0, 2, and 4. Maintenance dose: 200 mg every 2 weeks[1]			
Cimzia®*		£357.50	£ 9,326.92	£1,072.50[2]

Biologic	Dosing regimen	Cost per dose	Cost per year	Additional cost in Year 1
GLM				
Simponi®*	50 mg once per month, on the same date each month.[3]	£762.97	£9,155.64[4]	
IFX				
Remicade®*	3 mg/kg at Week 0, 2 and 4. Then 3 mg/kg every 8 weeks.[5]	£419.62 per vial (100 mg powder for concentrate for solution for infusion vials), 2 or 3 vials per administration	£5,747.48 (assuming no wastage), £8,210.69 (assuming full wastage)	£1,982.70
Inflectra® & Remsima®[6]		£377.66 (100 mg powder for concentrate for solution for infusion vials)	£5,172.76 (assuming no wastage), £7,389.66 (assuming full wastage)	£1,784.44
Flixabi®/ Renflexis®		£377.00 (100 mg powder for concentrate for solution for infusion vials)	£5,163.72 (assuming no wastage), £7,376.75 (assuming full wastage)	£1,781.33
Zessly®		NR		
Ixifi®		NR		

Key: NR: not reported; PAS: patient access scheme

Notes:

* Originator/ reference products

[1] Once clinical response is confirmed, 400 mg every four weeks may be considered

[2] Assuming no PAS arrangement

[3] Body weight up to 100 kg, 50 mg once per month, on the same date each month. Body weight exceeding 100 kg, initially 50 mg once a month (one the same date) for three to four doses, if treatment response is inadequate dose may be increased to 100 mg once a month.

[4] Based on standard dosing regimen for patient weighing less than 100 kg

[5] If treatment response is inadequate after 12 weeks, dose may be increased in 1.5 mg/kg increments every eight weeks to a maximum dose of 7.5 mg/kg every eight weeks. Alternatively, intervals between doses may be reduced, to a minimum dosing interval of 3 mg/kg every four weeks

[6] Cost per year was calculated assuming patient weight of 70 kg (as in TA375[50])

The annual costs of ADL, ETN, GLM and CTZ were estimated from the price of solution for injection *pre-filled pens* since these biologics are administered subcutaneously and can therefore be self-administered. Consultation with clinical experts confirmed that all the TNF inhibitors considered in this study except IFX are usually self-administered by people with RA at home.

Consistent with acquisition cost calculations in TA375,[50] the cost per annum of IFX was estimated using average weight of 70 kg.[8] IFX is administered intravenously (Section 4.1.9.1.7).

TA375[50] reports that the manufacturers of GLM provide the 100 mg dose at the same price as the 50 mg dose under a patient access scheme (PAS) arrangement, this discount does not affect the annual costs presented in Table 50 as these are based on the assumption that a patient weighs less than 100 kg.

Of note, the acquisition costs of the cheapest available pens for each drug are equivalent to the cost of the cheapest available dose, therefore annual acquisition costs for the self-administration route are equivalent to acquisition costs for biologics administered during outpatient visits.

The estimates for the additional acquisition costs for the first year (the last column of Table 50) are presented for information only. They were not used in any analyses since the population in this assessment are people experienced in biologics.

4.1.9.1.5 Dose tapering

According to EULAR recommendations for the management of RA with synthetic and bDMARDs,[93] tapering of bDMARDs should be considered in people in persistent remission after having tapered GC, especially if this treatment is combined with a conventional synthetic DMARD. In this context, tapering means reduction of dose (e.g. reducing etanercept 50 mg to 25 mg/ week (Smolen and colleagues, 2013[94]) or extension of interval between applications, 'spacing' (e.g. increasing the interval between ADL injections to 10 days rather than one week as in the Exeter biologic clinic recommendations described in Appendix 5).

The AG is aware that there is no gold standard on how dose tapering should be performed. Studies evaluating dose tapering have used different approaches. In clinical practice, dose tapering varies extensively depending on clinical opinion. For example, according to Exeter biologic clinic recommendations, when tapering the ADL dose, the dose should be reduced by one-third to 40 mg every three weeks and further reduced at three months to 40 mg every four weeks in people with LDA or remission (Appendix 5).

In the primary analysis, the assumption of reducing the dose by one-third (*the first dose reduction* in the Exeter biologic clinic recommendations, Appendix 5) was implemented (ADL dose tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).

- Dose is tapered in a proportion of people in each arm at the start of simulation.
- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored in all people on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), *and, therefore,* the administration cost is zero.
- The costs associated with flare management are:
- £423 per flare for diagnostic investigations
- £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.
- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.
    - The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
    - The duration of AE is 28 days.
    - The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 42), while the assumption of halving the dose (*the second dose reduction*) was explored in a sensitivity analysis (Table 73).

4.1.9.1.6 Wastage

The dose tapering strategy suggested in the Exeter biologic clinic recommendations (Appendix 5) is spacing. Therefore, when this tapering strategy is used, there is no wastage of the self-administered drugs due to partial use of the dose in pre-filled pens. Clinical advice (Dr Rich Haigh, Royal Devon and Exeter NHS Foundation Trust) indicated that wastage of IFX due to partial use of vials is usually avoided.

- In the primary analyses, however, wastage of £370 per patient-year was incorporated (ADL dose tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).
- Dose is tapered in a proportion of people in each arm at the start of simulation.

- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored *in all people* on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), and, therefore, the administration cost is zero.
- The costs associated with flare management are:
    - £423 per flare for diagnostic investigations
    - £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.
- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.
- The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
- The duration of AE is 28 days.
- The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 42). This estimate was based on a survey conducted at the Royal Devon and Exeter NHS Foundation Trust (Dr Rich Haigh, personal communication). It was derived from data from 119 people with RA treated with biologics, and included missed doses and over-supply (defined as a delivery of treatment even if >4 weeks supply was available at home). It was assumed that £370 per year was wasted, on average, in people on a full dose of a TNF inhibitor, while in people on a tapered dose, wastage was reduced proportionally to the reduction in treatment dose. In scenario analyses for other biologics, the treatment wastage was assumed to be proportionate to the drug acquisition price. The effect on the outcome of no wastage was explored in a sensitivity analysis (Table 73).

4.1.9.1.7 Drug administration costs

ADL, ETN, GLM, and CTZ are usually self-administered via subcutaneous injection using a pre-filled pen. In this scenario, there is no administration cost for delivery. Alternatively, these drugs may be administered by a district nurse. The average administration cost assumed in TA375[8] (which was based on an estimate reported in TA247[95]) was £2.61 (cost

year 2012). Since this cost is quite low and self-administration of the drugs listed above is very common, the effect of the assumption that subcutaneous administration of the drugs is performed by a nurse was not considered in the analyses reported here.

The administration cost for IFX is considerably higher since it is administered intravenously over a two-hour period. Patients are observed for at least one to two hours post-infusion for acute infusion-related reactions. Patients may be pre-treated with; e.g. an antihistamine, hydrocortisone and/or paracetamol and infusion rate may be slowed in order to decrease the risk of infusion-related reactions especially if infusion-related reactions have occurred previously.[96] Clinical advice indicated that IFX is typically administered in outpatient settings.

In the analysis for DG22[97] (Table 51), administration cost for IFX was estimated to be £287.93 per infusion; in a more recent technology appraisal, TA329, the administration cost was £297 per administration.[98]

Table 51: Administration cost of infliximab

Source	Cost per administration reported in the source	Cost year
TA329[98]	£297	2015
DG22[97]	£287.93	2014

Grant Smith (specialist pharmacist, Royal Devon & Exeter NHS Foundation Trust), advised us that in the NHS Foundation Trust, the cost of IFX administration is based on healthcare resource groups (HRGs) for Inflammatory Bowel Disease without Interventions, with Complications and Comorbidities (CC) scores depending on patient type. The relevant HRGs from the National Schedule of Reference Costs (Year 2017–18)[7] are shown in Table 52.

Table 52: HRG codes from the National Schedule of Reference Costs – Year 2017-18, day case

Currency code	Currency Description	Number of FCE's	National Average Unit Cost
FD02E	Inflammatory Bowel Disease without Interventions, with CC Score 5+	254	£317
FD02F	Inflammatory Bowel Disease without Interventions, with CC Score 3-4	1,496	£287
FD02G	Inflammatory Bowel Disease without Interventions, with CC Score 1-2	15,187	£282
FD02H	Inflammatory Bowel Disease without Interventions, with CC Score 0	81,985	£283

Key: FCE: finished consultant episodes
Source: National Schedule of Reference Costs, year 2017-18[7]

The weighted average administration cost of £283/administration estimated across the unit costs for the HRG codes presented in Table 52 was adopted in the scenario analysis considering people with RA treated with IFX (Table 75).

4.1.9.1.8 Cost of testing

The cost of concurrent testing of drug and antibody levels in patients treated with ADL and tested using Promonitor kits was estimated by Jani and colleagues (2016).[53] The study assumed the NHS perspective for identifying the resource use and cost per patient of providing anti-TNF inhibitor drug level and ADL antibody (ADAb) testing. Direct medical costs associated with providing the test were identified. The costs were determined from the point of a patient established on treatment (for ≥3 months) presenting to clinic, to the results being fed back to the clinician to inform a treatment decision (see Section 4.1.9.1.8 for further details on the cost components). In this study, it was assumed that most hospital laboratories would have the necessary room requirements and stock standard equipment required to perform ELISAs, and the following items of resource use were therefore excluded: equipment costs of centrifuge systems, ELISA readers, pipettes, personal protective equipment, phlebotomy equipment costs, overhead and capital costs. The mean cost *per patient per test* was £152.52 (2015 prices) if 40 samples were tested simultaneously. The pre-testing phase incurred the highest costs, which included booking an additional appointment to acquire trough blood samples, which was the key driver of costs per patient (67% of the total cost); labour accounted for 10% and consumables for 23% of the total costs. This study was an audit of practice in North West England. Refer to Table 53 and the following sections for further details on the cost components.

Table 53: Resource use and costs of implementing drug level and immunogenicity testing per patient in a hospital setting (Jani and colleagues, 2016)

Type of resource use	Cost
Phase 1: pre-testing	
Outpatient appointment for discussion about need for test	£2.35
Clerical staff (to book the appointment and send out a letter to patient)	£1.15
Appointment for trough blood levels	£102
Phase 2: analysis of samples	
Receipt and labelling of samples – central specimen reception	£2.22
Data entry of patient information to lab system	£2.22

Type of resource use	Cost
Sample preparation – extraction of serum from blood	£2.22
Transport, receipt and storage of sample – immunology lab	£2.22
Preparation of reagents (wash solution, setting up assay, conjugate)	£3.20
ELISA kit	£700
Pipette tips for ELISAs	£6.00
Semi-deep well plates for ELISAs	£2.20
Troughs for ELISAs	£1.00
Retrieval of patient/IQC samples from storage	£2.13
Checking and sorting samples to match worklist	£2.13
Pipetting samples onto ELISA plate	£4.26
Pipetting calibrators, IQC samples and incubation of samples	£2.13
Washing ELISA plate and addition of conjugate	£2.13
Washing ELISA plate and addition of substrate	£2.13
Addition of stop solution	£1.06
ELISA plate reading and printing of results	£2.13
Technical validation involving review of Internal quality control	£1.06
Results transcribed to worksheet	£1.06
Data entry of results to patient record in lab system	£2.13
Transcribed results/data entry reviewed by a second independent biomedical scientist	£1.06
Clinical authorisation using reference range/delta check failure results	£2.54
Hardcopy report sent to clinician	£2.11
Phase 3: Treatment decision	
Interpretation of results by rheumatologist	£3.92
Discussion with patient (phone call)	£3.47
Letter with results and decision	£2.16
Total costs (best case to worst case scenario)	£152.52 (£147.68-159.24)

Source: Table 1 in Jani and colleagues (2016)[53]

At the Exeter Clinical Laboratory (Royal Devon and Exeter NHS Foundation Trust) which conducts approximately 80% of testing for monitoring biologics in the UK, the fully recovered cost of TNF testing (staff, reagents, consumables, over-heads and depreciation of equipment) is [redacted] per test (Dr Timothy McDonald, personal communication); this includes testing of both drug and antibody levels and covers all the components in Phase 2 reported in Jani and colleagues (2015)[53] (Table 53). Of note, at the Exeter Clinical Laboratory only IDKmonitor test kits are currently used for clinical services.

Dr Timothy McDonald advised us that laboratories which conduct TNF testing have previously negotiated arrangements with the manufacturers of bDMARDS to cover the cost of biological monitoring, including assays and personnel costs. However, based on advice from Dr Meghna Jani, that might vary by geographical area and only for certain biologics (e.g. newer biosimilars).

4.1.9.1.9 Assay costs provided by the manufacturers

For the economic analysis, the cost of reflex and concurrent testing for each assay were derived from information request documents submitted by the manufacturers of the test kits.

Along with the list prices for the Promonitor test kits (Table 54 and Table 55), Grifols proposed price discounts (which depend on the uptake of testing) for test kits used in singlet or duplicate, concurrent or reflex testing with different number of tests per year. Therefore, additional cost-utility analyses were conducted for each level of discount and each type of testing. The resulting ICERs are not presented in the report due to the fact that they are very uncertain. However, the results are available in the model developed by the AG.

Grifols states in the request for information document that Promonitor only needs to be run once while the other ELISA-based tests are run in duplicate. However, our clinical expert Timothy McDonald advised us that the statement from Grifols is not correct – the other ELISA tests do not need to be analysed in duplicate (personal communication, December 2018).

The quality of information provided by manufacturers regarding the cost of testing, varied considerably. Only the manufacturer of Promonitor test kits (Grifols) and LISA-TRACKER (Cambridge Life Sciences) provided both the cost of reflex testing and concurrent testing (Table 54 and Table 55).

Data provided by the manufacturers regarding RIDASCREEN and IDK Monitor, reported only the costs of reflex testing. Furthermore, the information submitted regarding MabTrack

ELISA and Sanquin Diagnostic Services was impossible to interpret as some data appeared to be missing. Consequently, it was not possible to calculate the assay costs.

Dr Timothy McDonald advised us that the number of samples analysed per assay and, therefore, the cost per sample may vary in clinical practice. However, at large laboratories receiving a high number of referrals, it is likely that the maximum number of samples would be analysed. Therefore, in order to estimate the lower bound for the cost of testing, we assumed that the maximum number of samples per assay is analysed.

4.1.9.1.10 Processing costs

In addition to assay costs, the cost of testing also depends on processing costs, such as administration and laboratory personnel time. In the study conducted by Jani and colleagues (2016),[53] it was assumed that during the pre-testing phase (Phase 1, Table 53), one outpatient appointment with a consultant rheumatologist is required to discuss the need for testing, followed by an appointment with a phlebotomist or clinical support worker to obtain blood trough levels. Regarding the testing phase (Phase 2, Table 53), it was assumed that hospital laboratories would have the basic materials required to conduct ELISA, so several resource use items were excluded from the micro-costing exercise. This study reported that additional costs associated with laboratory personnel time processing the samples would be incurred during the testing phase (Table 53).

Finally, Jani and colleagues (2016)[53] reported that the treatment decision stage requires interpretation of results by a consultant rheumatologist, discussion of the results with patients via a telephone call, and lastly a letter outlining results and treatment decision (Phase 3, Table 53).

Table 54: Costs of reflex testing provided by the manufacturers of the test kits

Test		Singlet testing of patient samples			Duplicate testing of patient samples		
		Number of samples analysed per assay	Cost per assay	Cost per sample	Number of samples analysed per assay	Cost per assay	Cost per sample
IDK Monitor[a]	Drug levels monitoring	80	£855.00	£10.69	40	£855.00	£21.38
	Anti-drug antibody monitoring	90	£775.00	£8.61	45	£775.00	£17.22
Promonitor[b]	Drug level monitoring	80	£704	£8.80	40	£700	£17.50

Test		Singlet testing of patient samples			Duplicate testing of patient samples		
		Number of samples analysed per assay	Cost per assay	Cost per sample	Number of samples analysed per assay	Cost per assay	Cost per sample
	Anti-drug antibody monitoring	80	£704	£8.80	40	£700	£17.50
RIDASCREEN[a]	Drug level monitoring	96	£565.00	£5.89	48	£565.00	£11.77
	Anti-drug antibody monitoring	96	£775.00	£8.07	48	£775.00	£16.15
LISA-TRACKER[a]	Drug monitoring	48[d]	£836.77	£17.43	24	£836.77	£34.87
	Anti-drug antibody monitoring	48[d]	£836.77	£17.43	24	£836.77	£34.87
MabTrack	Drug monitoring		TBC			TBC	
	Anti-drug antibody monitoring		TBC			TBC	
Sanquin Diagnostics	Drug monitoring		TBC			TBC	
	Anti-drug antibody monitoring		TBC			TBC	

Key: ADM, Adalilumab; IFX, Infliximab; TBC: to be confirmed

Notes:

[a] Costs exclude VAT

[b] Cost inclusive of VAT

[c] Unclear whether cost includes or excludes VAT

[d] In the information request documents, the manufacturer provided only the cost of the 48-well assay, the cost of the 96-well assay was not provided as the manufacturer reported that this assay is rarely purchased.

Table 55: Costs of concurrent testing provided by the manufacturers of the test kits

	Singlet testing of patient samples			Duplicate testing of patient samples			Information source
Test	**N samples analysed per assay**	**Cost per assay**	**Cost per sample**	**N samples analysed per assay**	**Cost per assay**	**Cost per sample**	
Promonitor[a]	80	£700	£8.75	40	£700	£17.50	Request for information submitted by Grifols
LISA-TRACKERa	96	£1,550.77	£16.15	48	£1,550.77	£32.31	Request for information submitted by Cambridge Life Science

Key: N: number of
Notes
[a] Costs exclude VAT.

4.1.9.1.11 Cost of sample transport

One of the very minor cost components considered by Jani and colleagues (2016)[53] was "Transport, receipt and storage of sample" which was £2.22 (2015 prices) per batch of 40 samples (refer to Table 1 in Jani and colleagues 2016).[53]

Blood samples are received at the Exeter Clinical Laboratory (Royal Devon and Exeter NHS Foundation Trust) as small parcels via Royal Mail. Clinical advice from Dr Timothy McDonald indicated that it is extremely unlikely that samples would be sent to Sanquin Diagnostic Services in the Netherlands as the transportation cost would be higher compared to that within the UK. Postage costs are approximately £4 per parcel shipped within the UK and approximately £10 per parcel shipped to Sanquin Diagnostic Services.[17] Moreover, sending samples abroad would lead to a longer turnaround time and take expertise out of the NHS (Timothy McDonald, personal communication, December 2018).

Therefore, in all analyses relevant to MabTrack and Sanquin Diagnostic Services, the cost of sample transport of £10 was applied, while for all the other tests the postage of £4 per parcel was assumed (i.e. it was assumed that parcel would be posted to a laboratory within the UK).[17]

4.1.9.1.12 Frequency of testing

Rosas and colleagues (2015)[58] reported the total number of drug and anti-drug antibody monitoring tests in RA patients in remission over a two-year period (94 tests in 45 patients), which is approximately one test per patient per year (refer to Section 2.3 for further details on this study).

Dr Meghna Jani, confirmed that in people in remission/under routine follow up, TNF testing may be conducted once a year. However, if tapering is performed on the basis of drug level, a clinician would typically check drug levels at least every six-months to ensure that the level has not dropped too low.

In the primary analysis, one TNF test per patient-year was assumed, while six-monthly testing was modelled in a sensitivity analysis (Table 73).

4.1.9.1.13 Reflex versus concurrent testing

Dr Timothy McDonald (Exeter Clinical Laboratory, Royal Devon and Exeter NHS Foundation Trust), advised that TNF testing for blood and antibody levels is usually done concurrently. Blood samples sent to the Exeter Clinical Laboratory are kept frozen for one month, and the likelihood of performing antibody testing one month after testing trough levels is extremely low.

In this unlikely scenario when reflex testing is performed, an additional phlebotomy appointment (which is the key driver of the testing cost) would not be required (assuming that storage of blood samples is a common practice at test laboratories). Hence, the cost difference between reflex and concurrent testing would be defined by the proportion of patients with undetectable drug levels (for whom antibody testing would be requested) and the cost of phone calls to the laboratory to request antibody testing. However, there is no universal agreement on which drug levels should be considered undetectable (Dr Timothy McDonald, personal communication, December 2018).

To estimate the cost difference between reflex and concurrent testing, the proportion of people with low drug levels was derived from Chen and colleagues (2015).[65] The authors investigated the impact of ADL dose-halving on therapeutic responses and drug levels in people with RA. Serum ADL trough levels were determined at baseline and at Week 24 of dose-halving therapy using sandwich ELISA (Progenika Biopharma). The minimal detectable ADL levels were 0.002 mg/mL. In this study, three of 64 (4.7%) participants, who developed anti-adalimumab antibodies at Week 24 of dose-halving, had very low drug levels. In these participants, ADL trough levels markedly declined to very low levels (2.28, 1.92 and 2.21 mg/mL at baseline to 0.024, 0.024 and 0.004 mg/mL at Week 24 of dose-halving, respectively).

Laine and colleagues (2016),[2] reported low drug levels (<5 µg/mL) in 35.8% of people with RA, treated with ADL, from the clinical sample registry of United Medix Laboratories Ltd in Helsinki, Finland. All the samples included in the database had been sent to the laboratory on a clinical basis (i.e. none of the samples were from clinical studies). Drug levels were measured at Sanquin Diagnostic Services.

However, there is no universal agreement as to what to consider low drug level in people with RA treated with biologics (Dr Timothy McDonald, personal communication, December 2018).

The estimate of 35.8% for the proportion of people with low drug level was adopted in a scenario analysis for reflex testing.

In Jani and colleagues (2016),[53] a phone call to discuss a treatment decision with a patient was assumed to take, on average, 5.3 minutes at a cost of £3.47. Dr Timothy McDonald confirmed that this would also be a reasonable cost estimate for an additional phone call to the Laboratory to request additional testing on stored blood.

4.1.9.1.14 Singlet versus duplicate testing

The estimated costs of performing ELISAs once per patient *(singlet)* and *duplicate* reflex and concurrent testing using Promonitor are shown in Table 56.

Table 56: Costs of *singlet* and *duplicate* reflex and concurrent testing using Promonitor test kits

	Phlebotomy appointment		% of patients tested		Cost of phone call	Total cost, per patient	
	Initial	Additional	Trough level	Ab level		Duplicate	Singlet
Concurrent	Yes	No	100%	100%	NA	£161.73*	£141.66
Reflex	Yes	No	100%	4.7%[1]	£3.47	£146.15	£135.65
Reflex	Yes	No	100%	35.8%[2]	£3.47	£152.39	£138.77
Concurrent	No	No	100%	100%	NA	£56.30	£36.23
Reflex	No	No	100%	4.7%[1]	£3.47	£40.72	£30.22
Reflex	No	No	100%	35.8%[2]	£3.47	£46.96	£33.34

Key: NA: not applicable

Notes: All costs are in 2017-18 prices inclusive of the cost of sample transport of £4 (for a small parcel shipped within the UK).

* Cost assumed in the primary analysis

[1] Based on Chen and colleagues (2015)[65]

[2] Based on Laine and colleagues (2016)[2]

Singlet testing incurs a lower cost compared to *duplicate* testing. However, it is less precise. Therefore, *duplicate* testing was selected in the base-case analysis conducted by Jani and colleagues in the microsocting study (Jani and colleagues, 2016).[53] However, based on clinical advice, *singlet* testing is more common in the UK (Dr Timothy McDonald, personal communication, December 2018). Therefore, we adopted this approach in our primary analysis, and we conducted an additional analysis assuming *duplicate* testing.

In the primary analysis, the costs for reflex and concurrent testing using Promonitor test kits were based on the assumption that a phlebotomy appointment would be needed for

collecting the initial blood sample (Table 56). A scenario analysis assuming that samples are taken during an existing appointment was also conducted, as clinical advice indicated that this is common in clinical practice (Dr Timothy McDonald, Royal Devon and Exeter NHS Foundation Trust).

Of note, in all analyses, it was assumed that the maximum number of samples analysed per assay. Also, the test costs in Table 56 include postage of £4 per blood sample (assuming that it would be sent to a UK laboratory).

4.1.9.1.15 Training

Grifols stated in the request for information document that minimal additional training would be required before healthcare staff could use Promonitor safely and effectively. This was confirmed by clinical experts.

The company also wrote: "*All NHS laboratories will have experience with ELISA technology. The training for the Grifols Triturus Automated ELISA platform takes two days and Grifols provide full technical and service support for the duration of the contract. Grifols also provide on-site support and demonstrations for users when running the assays manually.*"

4.1.9.1.16 Cost of managing disease health states

Ucar and colleagues (2017)[11] and Arango and colleagues (2017)[15] provided results on the median duration of remission. However, none of the sources related to the INGEBIO study provided definitions of remission. In the primary analysis, we incorporated the effect on costs and health outcomes of different duration of remission in the intervention and control arms. Based on published literature, active disease in people with RA is more costly to manage compared to disease in people in remission/LDA. The major healthcare costs (apart from drug acquisition costs) in people with RA relate to joint replacement surgeries, hospital stays and doctor appointments.

There are a range of classification systems and scales that have been developed to measure and monitor disease activity in RA (as well as scales that are commonly used to measure other domains such as disability or activity level, such as the Health Assessment Questionnaire (HAQ).[42] Functional capacity measured with the HAQ was found to be the strongest predictor of costs (Kobelt and colleagues, 2005).[99] Therefore, direct medical costs for hospitalisations, joint replacements and the number of outpatient visits were included by HAQ-dependency, by assigning an annual cost to mutually exclusive HAQ intervals.

4.1.9.1.17 Resource utilisation in RA patients stratified by HAQ

Barbieri and colleagues (2005)[5] reported resource utilisation in people with RA treated with IFX, stratified by four HAQ bands (Table 57). These estimates were used by Barbieri and colleagues (2005)[5] to evaluate the costs of managing people with RA beyond the first year of therapy and were based on data from the Norfolk Arthritis Register (NOAR). The NOAR cohort includes 1,236 adults who had swelling of at least two joints that had persisted from more than four weeks. This study reported that on average, the number of outpatient visits, hospital days and the proportion undergoing joint replacement surgery increased substantially with HAQ score (Table 57).

Table 57: The average six-month resource utilisation stratified by HAQ status (the estimates were based on an average for five years follow-up)

	Hospital days[1]	Number of outpatient visits[1]	% of patients who had joint replacement[1]	Total costs, per year
HAQ 0	0.2	0.6	0.3	£3,474
0 < HAQ ≤ 1	0.5	1	0.8	£9,060[2]
1 < HAQ ≤ 2	1.2	1.5	2.3	£25,450[2]
2 < HAQ ≤ 3	5.1	2.1	4	£46,602

Key: HAQ: health assessment questionnaire; NA: not applicable

Notes:

[1] The estimates of resource use for different HAQ scores are based on the NOAR cohort reported in Barbieri and colleagues (2005).[5]

[2] The costs were estimated from the unit costs in Table 58.

Average costs of an inpatient day, outpatient appointment and joint replacement surgery derived from the relevant HRG codes from the National Schedule of Reference Costs - Year 2017-18[7] are shown in Table 58 (derivation of the cost of surgery is explained in the next section).

Table 58: Unit costs

	Cost	Source
Outpatient attendance rheumatology	£146	National Schedule of Reference Costs - Year 2017-18
Inpatient day	£413	National Schedule of Reference Costs - Year 2017-18 – elective inpatient excess bed day for inflammatory, Spine, Joint or Connective Tissue Disorders, with CC Score 0-2 (HD23J)
Joint replacement surgery	£5,222	National Schedule of Reference Costs - Year 2017-18, weighted average over currencies fo hip and knee procedures for non-trauma: HN12- HN14 and HN22- HN24

Source: National Schedule of Reference Costs - Year 2017-18[7]

Mean HAQ scores for different levels of disease activity (remission, LDA, moderate disease activity [MDA] and high disease activity [HAD]) in people with RA were estimated by Radner and colleagues (2014)[6] (Table 59): the mean HAQ score based on the Simple Disease Activity Index (SDAI) was 0.39, the mean HAQ score for LDA was 0.72, and the moderate disease activity (MDA) and high disease activity (HAD) was characterised by the mean HAQ of 1.24.

Table 59: HAQ scores for the states of disease activity according to the SDAI, CDAI and DAS28

Type of index	Remission		LDA		MDA/HDA	
	Mean	SD	Mean	SD	Mean	SD
SDAI	0.39	0.58	0.72	0.68	1.24	0.75
CDAI	0.38	0.56	0.75	0.70	1.23	0.74
DAS28	0.46	0.62	0.60	0.66	1.24	0.74

Key: CDAI: Clinical Disease Activity Index; DAS28: disease activity score in 28 joints; HAD: high disease activity; HAQ: health assessment questionnaire; LDA: low disease activity; MDA: moderate disease activity; SD: standard deviation; SDAI: the Simple Disease Activity Index

Using this classification and the cost estimates from Table 57, the costs for remission (the analysis for Ucar 2017) and active disease (Arango 2017) were estimated from the relevant distributions weighted by the annual health management costs for different HAQ scores. The costs of managing mixed health states (such as remission/LDA in the analysis for Arango 2017, and LDA/active disease for Ucar 2017) were based on joint distributions for the relevant health states weighted by the annual health management costs for the different HAQ bands (Appendix 9).

Cost of joint replacement surgery

The weighted average cost of joint replacement surgery estimated from HRGs relevant to hip and knee procedures for non-trauma *across all clinical codes* (HN12- HN14 ans HN22-HN24, respectively) is £5,222 per surgery (National Schedule of Reference Costs (year 2017-18[7]).

Burn at al. (2018)[100] estimated the hospital reimbursement for total knee replacement (TKR) and total hip replacement (THR) surgeries in the English NHS between 1997 and 2014. Primary TKR and THR were about £6,000 per surgery, while revision surgeries were about £8,000 per surgery in 2016/17 prices. These estimates were derived from the NHS primary care records on 21,128 people with osteoarthritis or RA were included in the analysis. The authors reported on the downward trends in the costs of TKR and THR.

The average cost of joint replacement surgery in people with RA from the Royal Devon & Exeter NHS Foundation Trust is £5,061.80 (Appendix 8). The estimate was based on 15 surgeries conducted in April 2017 - September 2018. Of note, this estimate is lower compared to those from the National Schedule of Reference Costs (Year 2017–18),[7] Table 58) and Burn at al. (2018).[100] This might reflect the trend in the cost of surgery reported by Burn and colleagues[100] However, the sample size was very low and therefore this estimate may not be representative of the average cost of surgery for the RA patient population in the UK (SE £5,153).

The average costs of managing remission and LDA/active disease health states (Table 57) derived from the average cost of joint replacement surgery based on the HRGs from the National Schedule of Reference Costs (Year 2017-18)[7] (£5,222 per joint replacement surgery, Table 58) were used in the primary analysis; a sensitivity analysis was conducted using the average cost of surgery reported in Appendix 8.

In the analyses reported here it was assumed, based on clinical advice, that surgery may be performed anywhere in the treatment pathway. The AG is aware however that older people are more likely to require surgery for RA.

4.1.9.1.18 Resource utilisation in RA patients stratified by DAS28 score

Barnabe and colleagues (2013)[101] investigated health service utilisation and costs associated with managing people with RA. This study was conducted in Canada and costs were estimated in 2008 Canadian dollars. In this study, patient costs were stratified by disease status: sustained remission was defined as DAS28 ≤2.6 for more than one year, while non-sustained remission was defined as DAS28 ≤2.6 for less than one year; non-

sustained low disease activity was defined as DAS28 >2.6 but ≤3.2 for less than one year, and persistent MDA or HDA was categorised as DAS28 >3.2.

This study reported hospital costs, emergency room costs, outpatient clinic costs, physician visit costs, and other outpatient costs in addition to annual crude mean costs, by disease status. The annual crude mean costs reported by Barnabe and colleagues (2013)[101] did not include drug costs. In this study, the mean (SD) age of participants was 55.1 (13.3) years, while the mean (SD) disease duration was 13.6 (9.5) years. The cost estimates reported in this study are presented in Table 60.

Beresniak and colleagues (2011) also investigated the costs of managing people with RA. This study stratified patients by disease status but did not consider duration of remission.[102] Participants were categorised as either in remission (DAS28 <2.6), LDA (DAS28 ≤3.2), or MDA to HDA (DAS28 >3.2). The authors reported direct costs for the first six months (excluding drug costs), and the direct costs incurred for each subsequent six month period (Table 60).

Table 60: Cost of managing disease (Barnabe and colleagues [2013] and Beresniak and colleagues [2011])

Barnabe and colleagues (2013)[101]	**Sustained remission (DAS28 ≤2.6 ≥1 year) (n=175)**	**Non-sustained remission (DAS28 ≤2.6 <1 year) (n=400)**	**Non-sustained LDA (DAS28 >2.6 but ≤3.2, <1 year) (n=138)**	**Persistent MDA or HDA (DAS28 >3.2) (n=338)**	**Study characteristics**	**Patient population and characteristics**
Hospital costs	738 (1699)	2249 (7450)	1665 (4157)	3423 (8561)	Country: Canada Currency: CAND Cost year: 2008 Reference period: Annual costs per patient	Population: People with RA treated with anti-TNF therapy enrolled in the Alberta Biologics Pharmacosurveillance Program (ABioPharm). Age, years mean (SD): 55.1 (13.3) Disease duration, years, mean (SD): 13.6(9.5)
Emergency room costs	196 (34)	240 (525)	235 (555)	459 (1008)		
Outpatient clinic costs	585 (822)	667 (948)	725 (1120)	823 (1614)		
Other outpatient costs	275 (473)	398 (614)	422 (714)	598 (1038)		
Physician visit costs	1337 (894)	1592 (1266)	1564 (1078)	2089 (1774)		
Total annual crude mean cost (2008 CAND)	3131 (3259)	5146 (9110)	4611 (6020)	7392 (11212)		
Beresniak and colleagues (2011)[102]	**Remission (DAS28 <2.6)**	**Not achieving remission**	**LDA state (DAS28 ≤3.2)**	**MDA to HDA (DAS28 >3.2)**	**Study characteristics[2]**	**Patient population and characteristics**
Direct costs for each subsequent 6-month period, excluding drug costs.	511 (162)	1,159 (339) per 6-month period	696 (240)	1,215 (405) per 6-month period	Country: France Currency: Euros Cost year: 2008 Reference period: Costs per patient per six month period	NA[1]

Key: CAND: Canadian dollars; DAS28: disease activity score in 28 joints; NA: not applicable; RA: rheumatoid arthritis; SD: standard deviation; TNF: tumour necrosis factor

Notes:

1. Data is hypothetical and is not based on a real patient sample, therefore no patient characteristics are reported.
2. This study estimated the resource use, stratified by disease activity states, using clinical guidelines, standard practice and exisiting evidence.

The average *annual* costs of managing remission and LDA/active disease estimated from those in Barnabe 2013 and Beresniak 2011 (in GBP 2017-18 prices) are presented in Table 61.

Table 61: Average annual costs of managing remission and LDA/active disease based on Barnabe and colleagues (2013) and Beresniak and colleagues (2011) (in GBP 2017-18 prices)

Source	Remission	LDA/active disease
Barnabe and colleagues (2013)[101]	£5,695[1]	£7,090[2]
Beresniak and colleagues (2011)[102]	£6,170[3]	£7,196[4]

Key: GBP: Great Britain Pounds; LDA: low disease activity

Notes:

[1] The cost for sustained remission from Barnabe 2013 was used to approximate the cost of managing remission.

[2] The esrtimate was calculated from the costs of non-sustained low disease activity and persistent, moderate or high disease activity reported in Barnabe 2013.

[3] Estimated from the cost of remission in Beresniak 2011.

[4] The estimate was calculated from the costs of LDA and MDA to HDA reported in Beresniak and colleagues (2011)[102]

The costs reported in Barnabe and colleagues (2013)[101] and Beresniak and colleagues (2011)[102] were first converted to pound sterling based on purchasing power parities (PPP) and inflated to 2017-18 prices using the healthcare price index (Section 4.1.9.1.1). Then the average cost of joint replacement surgery of £5,222 was added to all the estimates.

Importantly, the estimates of the costs of managing disease in people with RA in remission and LDA/active disease based on the French and Canadian studies (Table 61) are similar, but they differ substantially from those based on the HAQ bands (see Table 57 in the previous section). However, the data on costs used in Barnabe and colleagues (2013)[101] and Beresniak and colleagues (2011),[102] were from 2008, and therefore might not reflect the current clinical practice. Furthermore, these estimates are not directly relevant to the NHS; they were not used in any analyses and are presented here for information only to enable reader to better understand potential differences in clinical practice across the countries.

4.1.9.1.19 Cost of managing flares

The cost of managing flares is another important consideration that needs to be parameterised in the model. A study published by Maravic and colleagues (2005)[4] investigated the costs associated with managing flare-ups in people with RA. This study used a survey method to collect data regarding rheumatology practice for managing a hypothetical case of a flare-up, in an individual with a 10-year history of RA in a French setting.

A survey questionnaire was completed by 917 practicing rheumatologists. Over 80% of the respondents recommended measuring laboratory inflammation parameters, complete blood cell counts, liver enzymes, serum creatinine, and radiographs (hands, anteroposterior

cervical spine view, wrists, knees); 50–70% recommended additional cervical spine incidences, elbow and chest radiographs, and bone absorptiometry. Adding anti-TNF therapy (24%) or another DMARD (10%), increasing the methotrexate (MTX) dosage (24%), and substituting leflunomide for MTX were the main recommended treatments. Most respondents suggested continuing the GC in the same dosage (61%) or a higher dosage (36%). Analgesics and non-steroidal anti-inflammatory drugs were recommended by 65% and 41% of respondents and rehabilitation therapy by 83%.

This study focused on investigational costs and treatment costs; rheumatology appointments were not considered. Only *the total costs* of various types of tests and treatments were reported (Table 62).

Table 62: Cost of managing flares reported in Maravic and colleagues (2005) (2001 prices)

	Mean
Diagnostic investigations	
Laboratory tests	80
Other tests	276
Total 1	356
Treatment for one month	
DMARDs (n=884)	*724*
Glucocorticoids (n=901)	11
Analgesics (n=588)	17
Anti-inflammatory drugs (n=348)	14
Other treatments (n=130)	6
Total 2	746
Total cost 1+2	1,105

Key: DMARDs: disease modifying anti-rheumatic drugs
Note: Costs were based on dosages reported by respondents and brand names.
Source: Maravic and colleagues (2005)[4]

The total monthly cost of DMARDs (€724) was composed of the costs of MTX + ETN or IFX, MTX alone and other combinations with MTX. ETN and IFX are the major cost components contributing to the monthly cost because methotrexate is relatively inexpensive compared with the biologics. Since we modelled the cost of biologics separately, we did not include the cost of DMARDs in the cost of managing flares to avoid double-counting.

The cost of diagnostic investigations per flare (€356, Table 62) and the monthly cost of treatment excluding DMARDs (€24) were converted to pound sterling based on PPP and inflated to 2017-18 prices using the healthcare price index (Section 4.1.9.1.1) resulting in £423 and £68 for diagnostic investigations (per flare) and monthly treatment, respectively

(ADL dose tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).

Dose is tapered in a proportion of people in each arm at the start of simulation.

- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored in all people on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), *and, therefore,* the administration cost is zero.
- The costs associated with flare management are:
- £423 per flare for diagnostic investigations
- £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.
- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.
    - The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
    - The duration of AE is 28 days.
    - The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 42).

4.1.9.1.20 Cost of managing adverse events

In TA375, the weighted average cost of serious infection in RA patients was estimated to be £1,479 based on relevant NHS costs (NHS reference costs schedules 2010-11), weighted by inpatient activity (TA375 report, p. 311). Relevant HRG codes were identified based on Lekander and colleagues (2010). Conservatively the without complications and contraindications HRG costs were used.

The average cost inflated to 2017–18 prices using the healthcare price index (Section 4.1.9.1.1) was £1,622 (per infection). This cost was incorporated in our analysis (ADL dose

tapering is implemented by increasing the interval between doses from two to three weeks (i.e. by spacing doses).

Dose is tapered in a proportion of people in each arm at the start of simulation.

- Some people may flare after reducing the dose of their TNF inhibitors (Bykerk and colleagues, 2016).
- The full dose of ADL is restored in all people on tapered doses when they flare.
- Treatment wastage is £370 per patient-year in people on full dose; it is reduced proportionally to the reduction in treatment dose.
- ADL is self-administered (usually at home), *and, therefore,* the administration cost is zero.
- The costs associated with flare management are:
- £423 per flare for diagnostic investigations
- £68 per month for treatment (excluding the cost of DMARDs)
- The annual per-patient costs of managing remission and LDA/active disease health states are £11,409 and £18,889 respectively.
- The cost of managing an AE is £1,622.
- The utilities for remission and LDA/active disease health states are 0.717 and 0.586, respectively.
- The disutility of flare is 0.140.
- The duration of flare is seven days.
    - The rates of AEs in people on full and tapered doses are 3/100 and 2/100 patient-years, respectively.
    - The duration of AE is 28 days.
    - The time horizon is defined by the follow-up in Ucar and colleagues (2017).

Table 42).

4.1.9.2 Health related quality of life

A review of health-related quality-of-life (HRQoL) studies was conducted to inform the selection of utilities for the economic analysis. Utilities of remission/LDA and active disease health states, and disutilities for flares and serious adverse events (e.g. severe infections) identified in the review are described below.

4.1.9.2.1 Health state utility values

Ucar and colleagues (2017)[11] and Arango and colleagues (2017)[15] (INGEBIO) provided results on the average duration of remission in the intervention and control arms. However, none of the sources reporting INGEBIO study provided definitions of remission.

In Krieckaert and colleagues (2015),[1] health states were based on categorisation of DAS28 as below:

- remission (DAS28 <2.6)
- LDA (2.6≤DAS28<3.2)
- MDA (3.2≤DAS28≤5.1)
- HDA (DAS28>5.1).

The DAS28 is calculated from four components: tender, joint count, swollen joint count (performed by the clinician), visual analogue scale (VAS) score of the patient's global health and the laboratory parameter erythrocyte sedimentation rate (ESR). C-reactive protein (CRP) is more accurate as indicator of inflammation than ESR and it is also more sensitive to short-term changes. Disease activity states are defined in Table 63. The values are those reported in Bykerk and colleagues (2014)[16] using the DAS28-CRP score.

Table 63: Definition of disease activity types in Bykerk and colleagues (2014)

Type of disease activity	DAS28-CRP*
Severe	> 5.1
Moderate	3.2 – 5.1
Low	2.6 ≤ and < 3.2
Remission	< 2.6

Key: CRP: C-reactive protein; DAS28: disease activity score in 28 joints; ESR: erythrocyte sedimentation rate
Note: * DAS28-CRP is a modification of the DAS28 which includes the measured C-Reactive Protein (CRP) value, while the DAS28 uses the erythrocyte sedimentation rate (ESR) value instead.
Source: Bykerk and colleagues (2014)[16]

In the study conducted by Bartelds and colleagues (2011)[74] (Section 3.3.2.1), remission was defined as a DAS28 of less than 2.6 (at all consecutive measurements after a certain time point, with a minimum of two measurements of less than 2.6 for participants who discontinued treatment prematurely).

In Barnabe and colleagues (2013),[101] sustained remission was defined as DAS28 ≤2.6 for more than one year, while non-sustained remission was defined as DAS28 ≤2.6 for less than one year (Section 4.1.9.1.18).

In TA375,[50] non-responders, moderate responders and good responders were defined based on EULAR response criteria (see Table 64).

Table 64: The EULAR response criteria using the DAS28

DAS28 at endpoint	Improvement in DAS28 from baseline		
	≤ 1.2	> 0.6 and ≤ 1.2	≤ 0.6
≤3.2	good		
>3.2 and ≤5.1		moderate	
>5.1		none	

Key: DAS28: disease activity score in 28 joints
Source: Fransen and colleagues (2005)[49]

HSUVs estimated from HAQ according to SDAI, CDAI and DAS28

There are several composite scores to assess disease activity in RA. In this section, we consider the definitions of the disease states (i.e. remission, LDA, MDA and HDA) according to the SDAI, the Clinical Disease Activity Index (CDAI) and the DAS28 from Aletaha and colleagues (2007)[103] (Table 65).

Table 65: Cut-off points to separate the states of remission, and low, moderate and high disease activity using composite indices SDAI and CDAI and DAS28 score

Index	Remission	LDA	MDA/HDA
CDAI	≤2.8	≤10	≤22
SDAI	≤3.3	≤11	≤26
DAS28	<2.6	<3.2	<5.1

Key: CDAI: Clinical Disease Activity Index; DAS28: Disease Activity Score for 28 joints; HDA: high disease activity; LDA: low disease activity; MDA: moderate disease activity; SDAI: Simplified Disease Activity Index

Radner and colleagues (2014) collected data on clinical and laboratory characteristics (including CRP, ESR, number of swollen and tender joints, pain by VAS, patient global assessment of disease activity, evaluator global assessment of disease activity, and physical function by health assessment questionnaire [HAQ])), from 356 consecutive people with RA at routine clinic visits (every three to four months).[6] In total 716 visits were documented (median two clinic visits per person (range one to four).[6]

At baseline (according to the SDAI), 87 participants (24.4%) were in remission, 150 (42.1%) in LDA, 103 (28.9%) in MDA and 16 (4.5%) in HDA, but due to the low number of participants in the latter group, the last two groups were combined in further analysis.

The differences in functional disability envisaged by HAQ scores at three levels of disease activity (according to the SDAI) were evident, and similar conclusions were reached during a

sensitivity analysis, when the disease states were assessed according to the CDAI and DAS28 indices (Table 59).

The HAQ scores were mapped to the EQ-5D values using the same formula as in Section 4.1.9.2.2 and presented in Table 66.

Table 66: EQ-5D scores for the states of disease activity according to the SDAI, CDAI and DAS28

Type of index	Remission		LDA		MDA/HDA	
	Mean	Range	Mean	Range	Mean	SD
SDAI	0.717	(0.56, 0.80)	0.634	(0.43, 0.80)	0.483	(0.22, 0.69)
CDAI	0.72	(0.57, 0.80)	0.63	(0.42, 0.79)	0.49	(0.23, 0.69)
DAS28	0.70	(0.53, 0.80)	0.67	(0.48, 0.80)	0.48	(0.23, 0.69)

Key: CDAI: Clinical Disease Activity Index; DAS28: Disease Activity Score for 28 joints; HDA: high disease activity; LDA: low disease activity; MDA: moderate disease activity; SDAI: Simplified Disease Activity Index

The utility values of 0.70 and 0.586 for remission and LDA/active disease health states, respectively, were used in the primary analysis, while HSUVs obtained from HAQ scores reported in TA375[50] (as described in the next section) were assumed in a scenario analysis.

The AG is aware of several algorithms for converting the HAQ score to utility in RA and that the estimates of utilities may vary when different mapping algorithms are used.[104] To address this uncertainty, HSUVs were estimated using a quadratic equation proposed by Malottki and colleagues (2011) and used in TA375 to map HAQ to EQ-5D scores.[14,50] In TA375 a comparison of published relationships between utility and HAQ was conducted.[50] Three of the eight studies in the comparison in TA375 reported data from the UK. Of the three studies, Bansback and colleagues (2007)[105]b included data for UK and Canadian patients and Kobelt and colleagues (2002)[106] included data for UK and Swedish patients and were therefore not considered relevant for the purposes of this analysis. Hurst and colleagues (1997)[107] included only people with RA in Scotland. Malottki and colleagues (2011)[14] used the data set from Hurst and colleagues (1997)[107] to estimate the coefficients of their mapping equation and therefore there is little difference between the two sources.

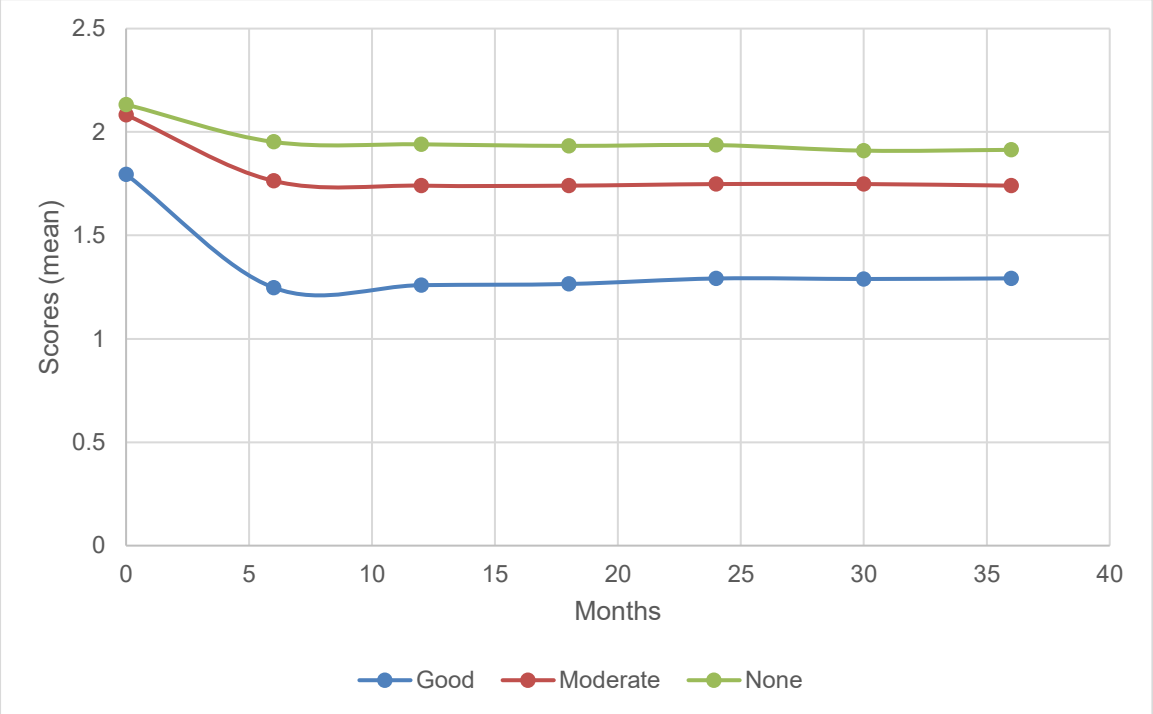
HSUVs estimated from HAQ by EULAR response category

In TA375,[50] the model was based on EULAR response category (good/moderate/none) for consistency with NICE guidance on biologics in RA and to align more closely to UK clinical practice in terms of the assessment of response to therapies.

The HAQ scores were estimated from the BSRBR-RA database (which contains values measured at six-month intervals for up to three years for all people with RA on the register), restricted to those with full set of baseline characteristics and at least two additional HAQ measurements while on bDMARDs.[50] The database included data from 10,186 patients. Of these, 2,417, 5,492 and 2,277 were classed as EULAR good responders, moderate responders and non-responders, respectively (Table 64).

Figure 10 shows the HAQ trajectory in people with RA treated with bDMARDs. It was observed that the mean HAQ scores for patients with good, moderate or no response (according to EULAR response criteria shown in Table 64) decrease during the first six month since the start of biologic therapy (where the scale of decrease grows with the level of EULAR response), then stabilise at around six months and remain quite flat over the remaining 2.5 years of measurement (Figure 10).

Figure 10: Mean HAQ by EULAR response category for patients receiving biologics



Source: Reproduced from Figure 94, in TA375,[50] p. 366)

The HAQ scores after six months of the biologic therapy for all three categories of responders were mapped to EQ-5D utilities using the same approach as that described in Section 4.1.9.2.2 , which elicited values shown in Table 67.

Table 67: Utility values based on BSRBR data

Type of patients	Number of patients in the BSRBR-RA dataset	HAQ score	Utility
Non-responders	2,277	1.95	0.237
Moderate responders	5,492	1.7	0.329
Good responders	2,417	1.2	0.496

Key: BSRBR-RA: British Society for Rheumatology Biologics Register in Rheumatoid Arthritis; HAQ: health assessment questionnaire
Source: TA375[50]

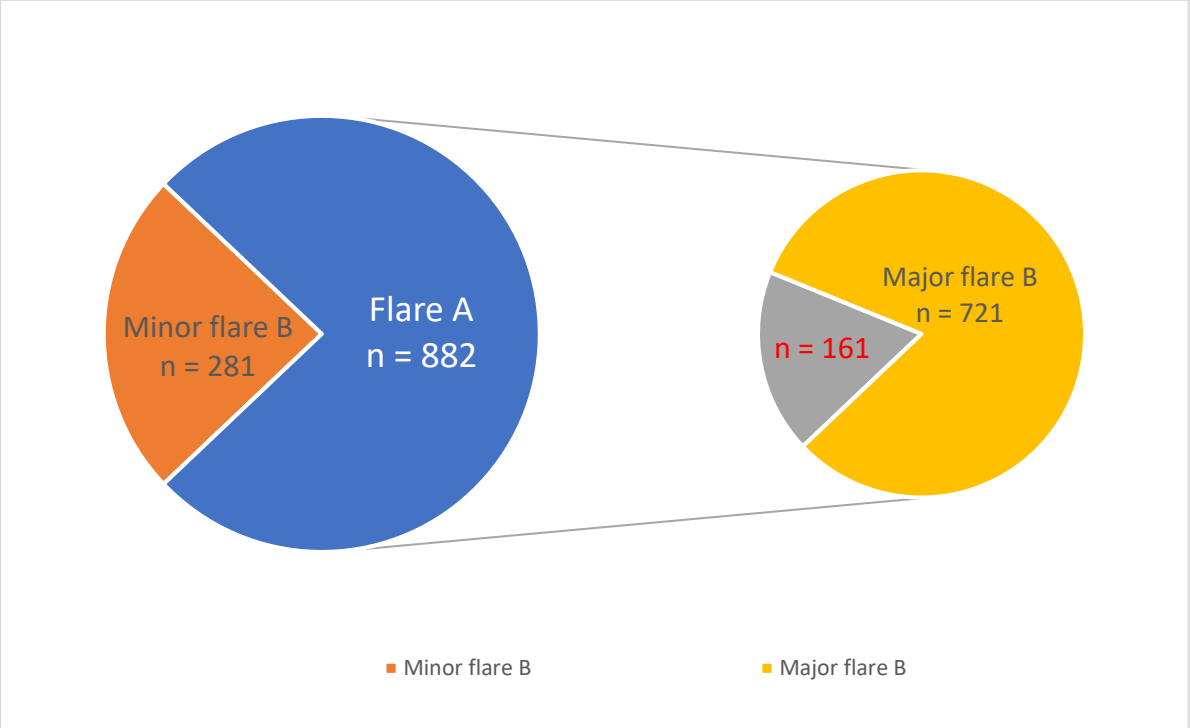
Hernández Alava and colleagues (2013)[108] argued that pain should be included as an explanatory variable when estimating QALYs from HAQ scores in people with RA. This approach was used in TA375.[50] However, the estimates presented in Table 67 were obtained without pain scores because the AG did not have access to patient-level data.

The utility for the remission health state was based on the utility value for good responders (0.496 in Table 67), while the utility for the LDA/active disease health state was estimated as a weighted average of the utility values for moderate responders and non-responders, resulting in the utility value of 0.302. These HSUVs were used in a sensitivity analysis.

4.1.9.2.2 Disutility of flare

The values of utility losses due to flares were obtained from the Dutch multi-centre, clinical study 'BeSt' which involved 508 participants treated-to-target for 10 years to achieve disease activity score (DAS28) of at most 2.4.[9] Since the concept of flare is not yet well-defined and no generally-accepted measure of its severity currently exists,[3,109] the BeSt study considers three types of flares named as 'A', 'minor B' and 'major B' (where the latter is a sub-category of the first) whose number of occurrences (observed during a total of 11,485 visits of all patients to a rheumatologist) is shown in Figure 11 and whose definitions, frequencies and HAQ scores are described in Table 68 (sufficient follow-up data were available only for 480 patients).

Figure 11: Three types of flares considered in Markusse and colleagues (2015)



Source: Markusse and colleagues (2015)[9]

Table 68: The definition of flares and their Health Assessment Questionnaire values

Type of flare	DAS28		Frequency		HAQ	
	current	previous	increase		Mean	SD
A	>2.4	any	≥0.6	321/480 (67%)	1.04	0.63
Minor B	>2.4	≤2.4	<0.6	159/480 (33%)	0.85	0.55
Major B *	>2.4	≤2.4	≥0.6	304/480 (63%)	0.96	0.60

Key: DAS28: disease activity score in 28 joints; HAQ: health assessment questionnaire; SD: standard deviation
Note:
* Major B is a subcase of A

Functional mobility of patients with respective types of flares was measured using the HAQ values (whose mean and standard deviation are also included in the table), which were mapped to the EQ-5D scores according to a quadratic equation

$$EQ5D = a - b_1 HAQ - b_2 HAQ^2$$

where coefficients a = 0.804, b_1 = 0.203 and b_2 = 0.045 were estimated from the UK data.[14] As this formula may return negative values for some high HAQ scores, which is sometimes regarded as controversial, it is recommended to adjust the values to zero. The loss of QALYs was computed as the difference between the utility values in the respective types of

flares and in the absence of flares (HAQ of 0.53). The estimated utility values are shown in Table 69.

Table 69: The definition of flares and their utility values

Type of flare	Utility			HAQ		Disutility
	mean - SD	mean	mean + SD	mean	SD	
A	0.339	0.544	0.713	1.04	0.63	-0.140
Minor B	0.432	0.599	0.739	0.85	0.55	-0.085
Major B *	0.378	0.568	0.725	0.96	0.60	-0.116

Key: HAQ: health assessment questionnaire; SD: standard deviation

4.1.9.2.3 Disutility of serious adverse events

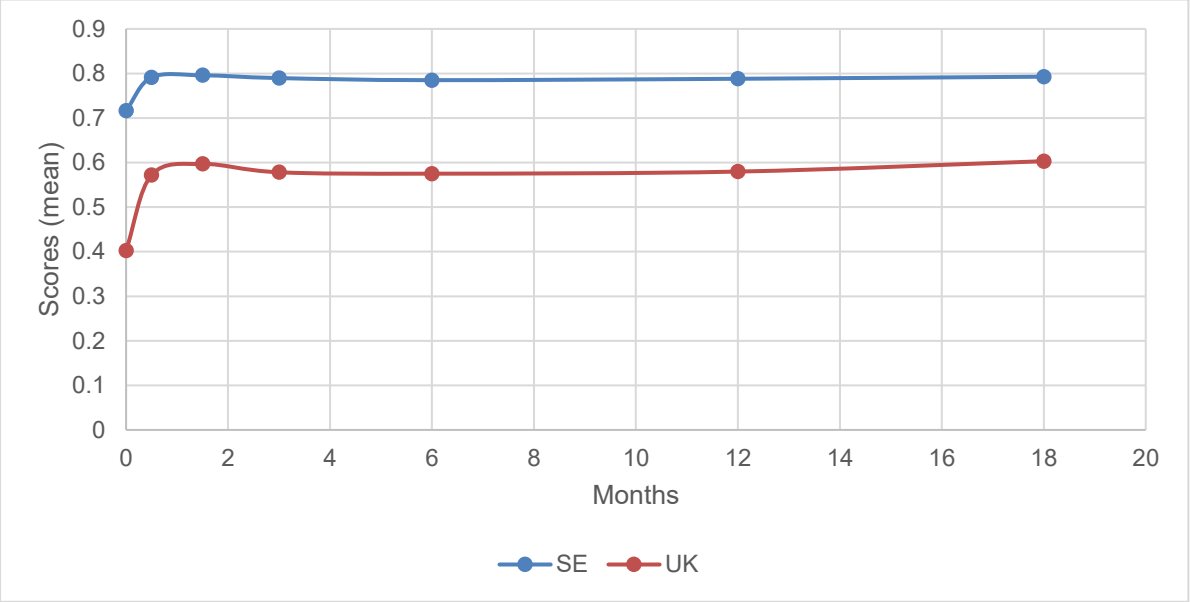
People with RA have increased susceptibility to serious infections due to features of RA, comorbidity and immunosuppressive treatment.[110] TNF-α inhibitors increase the risk of serious infection up to two-fold.[111]

A scenario analysis including serious adverse events was performed. The disutility value for England of 0.156 over four weeks (equivalent to the loss of QALYs of 0.012) associated with severe infections was estimated using the EuroQol's measure EQ-5D reported in the observational study "Genomics to combat Resistance against Antibiotics in Community-acquired lower respiratory tract infections (LRTI) in Europe" (GRACE) of the management of patients with acute cough/LRTI in primary care.[10] Data were collected in 13 European countries (including England and Wales) from adults (aged 18 years-plus) who reported to their primary care clinician with cough and LRTI.[10] The EQ-5D index scores were generated using the country-specific UK value set (the original data were collected from non-institutionalised adults in England, Scotland and Wales between August and December 1993 with a total of 2,997 participants).

4.1.9.2.4 Consistency between utility values

The observed discrepancies between utility values calculated in different countries may be due to differences in distinct preference sets for those countries. Based on data reported in Gülfe and colleagues (2016),[112] Figure 12 shows the discrepancies between EQ-5D scores obtained using British and Swedish preference sets for people with established RA being treated with TNF-inhibitors.

Figure 12: EQ-5D utility according to British (UK) and Swedish (SD) preference sets of patients with established rheumatoid arthritis treated with TNF-inhibitors



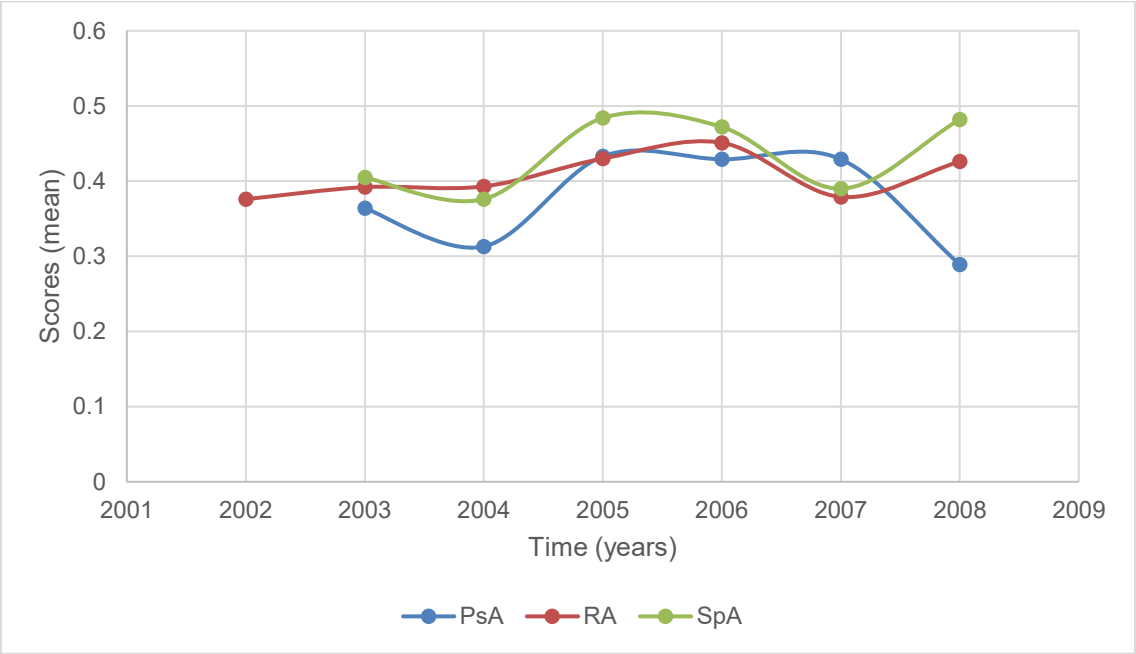
Source: Gülfe and colleagues (2016)[112]

It should be noted that in the analyses considering the duration of remission/LDA together with flares or AEs, there is a risk of double-counting of the effect of flares and AEs on HRQoL since it is possible that the disutilities have already been incorporated into the mapping equation from HAQ to utility. This is a limitation of these analyses.

In the INGEBIO study, a mixed disease population of patients was considered. This trial recruited 169 people including those with RA (n=63), PsA (n=54) and ankylosing spondylitis (AS) (n=52) (Section 2.3.2.1). A similar mixed disease population was condidered in Gülfe and colleagues (2010).[113] One of the aims of this study, was to analyse trends in baseline health utilities in people diagnosed with three types of arthritis (2,554 with RA, 574 with PsA and 586 with spondylarthritis [SpA]) who started treatment with TNF-α inhibitors, to address changes of utility during treatment and to understand the influence of previous courses of treatment. Data for the period from May 2002 to December 2008 were provided from the Swedish Arthritis Treatment Group (SSATG) register, which was set up in 2002 and has been collecting health utility data from routine clinical follow-up (time points of 0, 0.5, 1.5, 3, 6, 12, 24, 30, 36 months), and treatment courses were classified as either first, second, or third or more anti-TNF. Among three sub-populations, people with RA were characterised by older age, had tried more DMARDs, were more often treated with a concomitant DMARD, and were more often female compared to the other populations.[113]

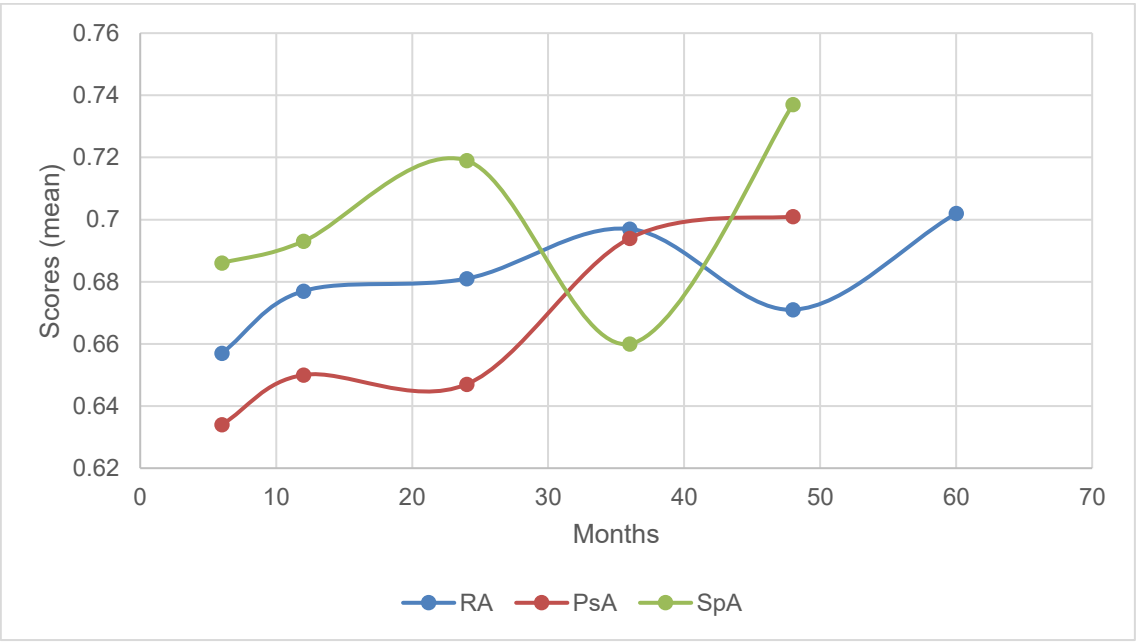
Figure 13 shows trends of baseline utility values at the start of treatment with TNF-α inhibitors for different diagnoses, and Figure 14 shows different response patterns in people with RA and those with PsA, and SpA.

Figure 13: EQ-5D at first initiation of treatment with TNF-inhibitors for rheumatoid arthritis, psoriatic arthritis, and spondyloarthritis during 2002-08



Key: PsA: psoriatic arthritis; RA: rheumatoid arthritis; SpA; spondyloarthritis
Source: Gülfe and colleagues, 2010

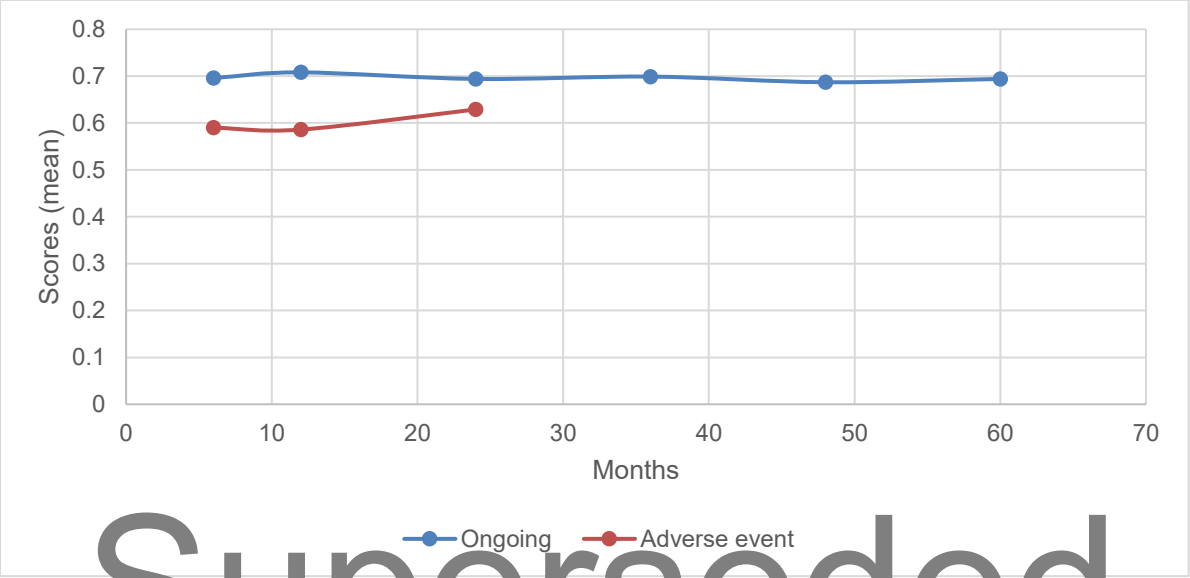
Figure 14: EQ-5D during the first courseof treatment with TNF-inhibitors in 2002-08 for people with rheumatoid arthritis, psoriatic arthritis, and spondyloarthritis



Key: PsA: psoriatic arthritis; RA: rheumatoid arthritis; SpA; spondyloarthritis
Source: Gülfe and colleagues, 2010

People with RA also demonstrated lower utility gain upon termination of the therapy independently of the reason for withdrawal of treatment compared to the other sub-populations. Figure 15 provides another possible way of estimating the disutility of adverse events as the averaged difference between the plotted values for ongoing treatment and those for withdrawal due to an AE.

Figure 15: EQ-5D during follow-up upon withdrawal from treatment

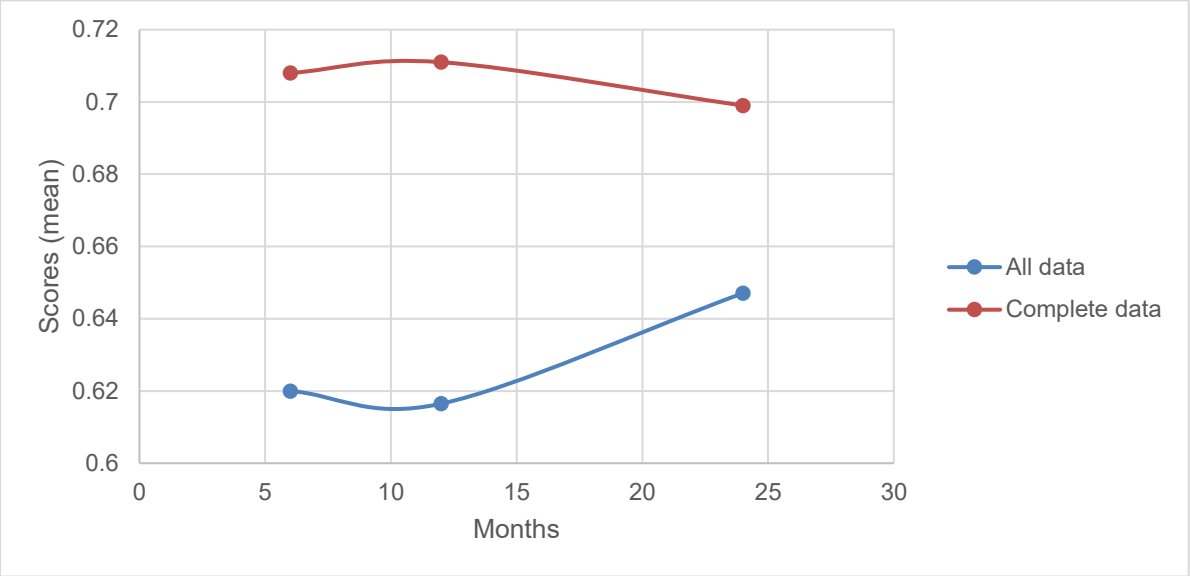


Key: PsA: psoriatic arthritis; RA: rheumatoid arthritis; SpA: spondyloarthritis

Source: Goeb and colleagues, 2010

Since it is difficult to obtain complete data in the observational setting. While using all available data increases the generalisability of the study, it also leads to lower improvement estimates as compared to using data only for those participants for whom complete follow-up information is available from all visits (incomplete data sets may be caused by withdrawal from treatment, for example) (Figure 16).

Figure 16: EQ-5D for people with rheumatoid arthritis (all participants versus participants with complete data) during the first and second lines of anti-TNF-α treatment



Key: PsA: psoriatic arthritis; RA: rheumatoid arthritis; SpA; spondyloarthritis
Source: Gülfe and colleagues, 2010

4.1.9.2.5 Mortality

Whilst there is evidence of an association between HAQ improvement and reduced mortality risk, the impact of TNF testing on mortality was not considered due to the short-term time horizon adopted in this study and a relatively small difference in the mean duration of remission/LDA across the treatment arms in the INGEBIO study.

4.1.10 Checking the model for wiring errors

The Excel code was checked in the following ways: all calculations were performed by one person and checked by another person, and the reasonableness of outputs given extreme input values was checked.

4.2 Cost effectiveness results

4.2.1 Adalimumab and Promonitor

4.2.1.1 Threshold analysis

The results of the threshold analysis, assuming the Promonitor test kit is used to monitor people with RA in remission/LDA recieving originator ADL (Humira®) are presented in Table 70 and Figure 17. Figure 17 shows the annual cost of ELISA-based testing at which TDM would become cost-effective at the two WTP thresholds used in NICE decision making for the range of ADL acquisition costs of £1,000–£9,180. Since the data reported in Arango and colleagues (2017)[15] are for a longer follow-up than that reported in Ucar and colleagues

(2017)[11], the results using the two different reports of the outcomes of the INGEBIO study are presented.

If the results of Ucar and colleagues (2017)[11] are used, then with the current price of originator ADL, testing would need to be cheaper than £391 per year in order for TDM to be judged as cost-effective. Using the the results presented in Arango and colleagues (2017);[15] however, there would be no cost of testing at which testing becomes cost-effective (because using these outcomes testing was estimated to be both more costly and less effective than standard care).

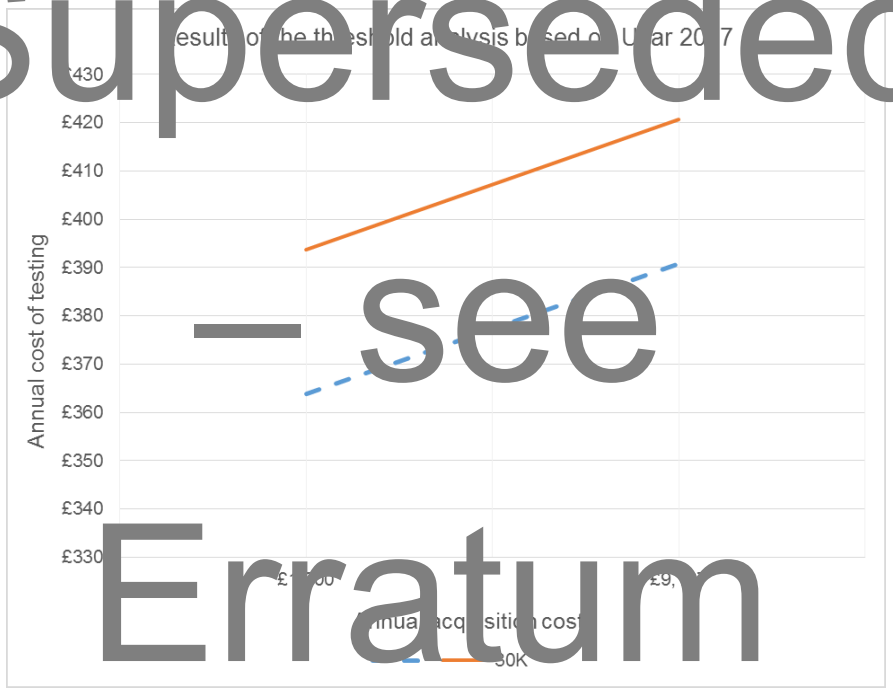
Table 70: Threshold value for the cost of testing at which NMB is zero

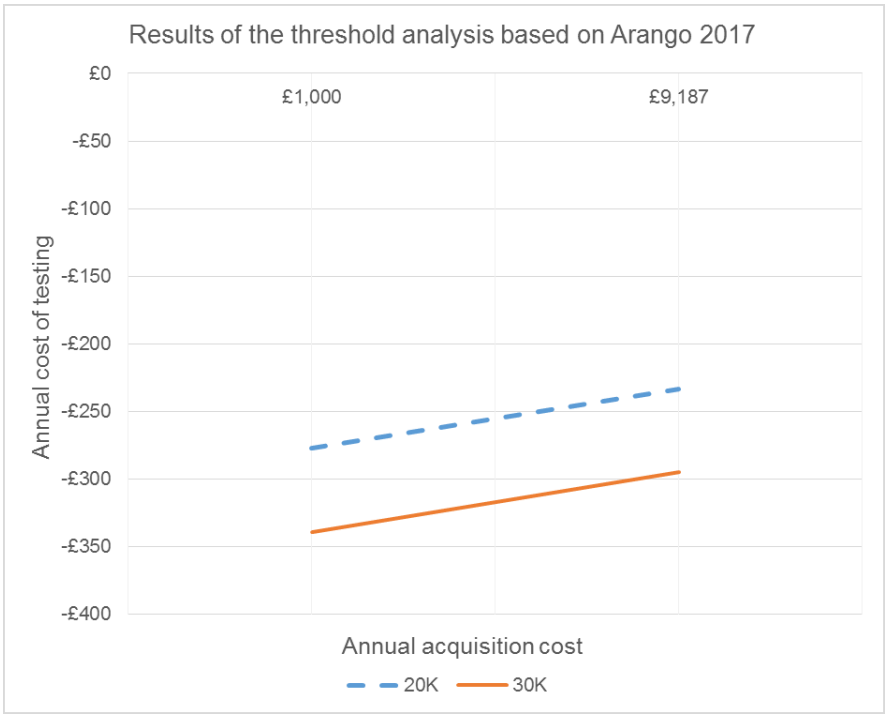
ICER threshold	Results based on INGEBIO study, Ucar and colleagues 2017	Results based on INGEBIO study, Arango and colleagues 2017
£20,000	£391	-£233
£30,000	£421	-£295

Key: ICER: incremental cost-effectiveness ratio; NMB, net monetary benefit
Source: Ucar and colleagues (2017)[11] and Arango and colleagues (2017)[15]

Such differences in the results are due to differences in the mean duration of remission (as reported in Ucar 2017) and remission/LDA (Arango 2017) between the control and intervention arms. Arango reported a longer duration of remission/LDA in the control group than in the intervention group (475.2 versus 460.2 days), while Ucar and colleagues 2017 reported a longer duration in the intervention group (344 versus 329 days in the control group).

Figure 17: Results of the threshold analyses using Ucar and colleagues (2017) and Arango and colleagues (2017)





Source: Ucar and colleagues (2017)[11] and Arango and colleagues (2017)[15]

These results are inconclusive for two reasons. First, because they are in opposite directions and, second because they are based on very small and uncertain differences in outcomes (QALY differences of less than 0.01). The negative value of the cost of testing at which NMB equal zero means that, when using the trial results as presented in Arango and

colleagues (2017),[15] there are no (positive) values of the cost of testing at which it would be a cost-effective option.

4.2.1.2 Cost-utility analysis

The incremental QALYs and incremental costs for testing versus standard care strategy are shown in Table 71, assuming:

- regular testing is undertaken in people with RA in remission/LDA treated with Humira® and tested using Promonitor
- the costs of testing are as in Jani and colleagues (2016)
- the frequency of testing is *one test per patient-year* and
- that testing of drug and antibody levels is done *concurrently (singlet dilution) in a UK laboratory.*

The outcome data were derived from two reports of the INGEBIO study, Ucar 2017 and Arango 2017.

As can be seen from Table 71, the main cost components are drug acquisition and the costs of managing health states. The main differences in costs between the intervention and control arms are the costs of managing health states and flares, and the cost of phlebotomy appointment. The main QALY components are those for the health states. The differences in QALYs for flares and AEs between the intervention and control arms are very small.

Table 71: Cost-effectiveness results in patients in remission/LDA treated with Humira® and tested using Promonitor

	Intervention arm	Control arm	Intervention vs. control
Based on Ucar and colleagues (2017)			
Costs			
Drug acquisition	£12,078	£12,120	-£42
Drug admin	£0	£0	£0
Drug wastage	£486	£488	-£2
Cost of managing health states	£19,071	£19,379	-£307
Cost of flare management	£281	£388	-£107
Cost of managing AEs	£64	£64	£0
Cost of phlebotomy appointment	£162	£0	£162
Other costs of testing	£30	£0	£30
Cost of sample transport	£6	£0	£6
Total costs (mean)	£32,178	£32,438	-£260
QALYs			

	Intervention arm	Control arm	Intervention vs. control
Remission	0.675	0.646	0.029
LDA/active disease	0.258	0.282	-0.024
Flares	0.002	0.002	-0.001
AEs	0.172	0.173	-0.001
Total QALYs (mean)	1.108	1.103	0.004
ICER (Cost / QALY gained)			ICER not relevant - Intervention dominates standard care
Based on Arango and colleagues (2017)			
Costs			
Drug acquisition	£13,075	£13,149	-£74
Drug admin	£0	£0	£0
Drug wastage	£527	£530	-£3
Cost of managing health states	£22,112	£21,757	£355
Cost of flare management	£303	£418	-£115
Cost of managing AEs	£69	£70	£0
Cost of phlebotomy appointment	£162	£0	£162
Other costs of testing	£30	£0	£30
Cost of sample transport	£6	£0	£6
Total costs (mean)	£36,284	£35,923	£361
QALYs			
Remission/LDA	0.838	0.865	-0.027
Active disease	0.112	0.092	0.020
Flares	0.002	0.003	-0.001
AEs	0.187	0.188	-0.001
Total QALYs (mean)	1.138	1.147	-0.009
ICER (Cost / QALY gained)			ICER not relevant - Standard care dominates Intervention

Key: ICER: incremental cost effectiveness ratio; QALY: quality-adjusted life year; vs: versus
Note: The postage was £4 per parcel

As with the threshold analyses, these results are inconclusive. First, because they are in opposite directions and, second because they are based on very small and uncertain differences in outcomes (QALY differences of less than 0.01). Furthermore, it is not possible to argue that either the analysis based of Ucar and colleagues (2017)[11] or that based on Arango and colleagues (2017) is more valid than the other – they both have significant weaknesses (refer to Section 2). The follow-up in Arango and colleagues (2017)[15] is over a longer time horizon (545 days in the control arm) that Ucar and colleagues (2017) (505 days

in the control arm).[11] The fact that these different analyses from the same study produce opposite estimates of effects and costs further highlights the uncertainty, which the economic analysis for this appraisal may only serve to amplify.

4.2.1.3 Sensitivity analyses

A number of sensitivity analyses were undertaken to explore the impact of parametric and structural uncertainty on the outcomes reported in Table 71 as follows:

- the impact of study follow-up on the outcomes of the economic analysis
- the percentage of people in whom the biologic was tapered
- the tapering strategy of dose halving
- the differential flare rate
- the duration of flare
- the disutility of flare
- the proportion of patients with flares in whom increase in medication dose is implemented
- the effect on the results of the uncertainty in the costs of managing remission and LDA/active disease states
- the health state utility values
- the effect on the results of flares only, i.e. when health states and AEs are not considered (as in Scenario C, Gavan, 2017)
- the total cost of treatment wastage (it was assumed to be zero)
- the effect of excluding the cost of the initial phlebotomy appointment
- the effect of testing in duplicates
- the effect of reflex testing for two assumptions on the proportion of people with low drug level: 4.7% (the lower bound) and 35.8% (the upper bound)
- the frequency of testing of two tests per year

These sensitivity analyses are detailed below and the results are shown in Table 73.

4.2.1.3.1 Treatment wastage

The assumption of no treatment wastage was explored in a sensitivity analysis, and had no impoact on the results, see Table 73.

4.2.1.3.2 Flare duration

A sensitivity analysis was conducted to evaluate the impact of increasing the duration of flare from 7 days (as assumed in Section 4.2.1.2) to 19 days (based on the weighted average

duration of flare derived from Bykerk and colleagues (2014)[16]and expert advice). Increasing the estimated duration of flare did not affect the results, see Table 73.

4.2.1.3.3 Proportion of flared patients in whom increase in medication dose was implemented

A US study conducted by Bykerk and colleagues (2014) reports statistics on flares management which shows that at least 45% of treatment strategies for coping with flares did not involve dose increase or any other change of medication (Bykerk and colleagues, 2014).[16]

Dr Rich Haigh (clinical advisor), confirmed that in about two-thirds of all flared patients on tapered doses the dose would be switched back to full.

We evaluated the effect of this assumption on the model results by assuming that 45% of people with flares would stay on the same (tapered) dose. It was also assumed that 100% of flared patients would stay on the same (tapered) dose.[114,115] Neither assumption impacted on the results, see Table 73.

4.2.1.3.4 Health state utility values

In this scenario analysis, utility values for remission and LDA/active disease health states were as shown in Table 72. Further details on how those values were derived are provided in Section 4.1.9.2.1. The result of the sensitivity analysis is shown in Table 73.

Table 72: Health state utility values

Health state	Utility value	Source
Remission	0.496	(Section 4.1.9.2.1)
LDA/active disease	0.302	(Section 4.1.9.2.1)

4.2.1.3.5 Impact of the cost of initial phlebotomy appointment

A scenario analysis assuming that trough samples are taken at the time of an existing appointment was also conducted, as clinical advice (Dr Timothy McDonald) indicated that this is quite common in clinical practice. The costs for reflex and concurrent testing for scenario analyses shown in Table 56 were assumed. Changing this assumption had no impact on the results, see Table 73.

Table 73: Sensitivity analyses (people in remission/low disease activity)

Sensitivity analysis	Assumptions	ICER		Source (relevant sections)
		Ucar and colleagues (2017)	**Arango and colleagues (2017)**	
Impact of flares only (health states and AEs are not included)	Only flares contribute to differential costs and QALYs	ICER not relevant - Standard care dominates Intervention	ICER not relevant - Standard care dominates Intervention	Scenario C (people in remission, Gavan 2017, Section 3.3.2.3**Error! R eference source not found.**)
Tapering strategy	Spacing: reduction of ADA dose to 40mg every 4 weeks	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	2nd dose reduction Exeter biologic clinic recommendations (Appendix 5)
Treatment wastage	No wastage	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Assumption
Flare duration, days	19	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Weighted average based on Bykerk and colleagues (2014)[16] and clinical advice
Proportion of flared patients in whom full dose is restored	55%	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Bykerk and colleagues (2014)[16] and clinical advice
	0%	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Clinical advice
Utilities[2]				
Remission	0.496	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Estimated from HAQ scores reported in TA375[50] (Fig. 94, p.366) (Section 4.1.9.2.1)
LDA/active disease	0.302			

Sensitivity analysis	Assumptions	ICER		Source (relevant sections)
		Ucar and colleagues (2017)	**Arango and colleagues (2017)**	
Disutility of flare	0.085	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Minor B type of utility (Table 69, Section 4.1.9.2.2)
	0.116	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Major B type of utility, (Table 69, Section 4.1.9.2.2)
Frequency of testing (tests/year)	2	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention	Rosas and colleagues 2015,[58]r clinical advice (Section 4.1.9.1.20)
Cost of testing	Refer to Table 56 for the cost of testing			
Duplicate concurrent testing with initial phlebotomy appointment		In all analyses, ICER not relevant -Intervention dominates standard care	In all analyses, ICER not relevant - Standard care dominates Intervention	Jani and colleagues (2015)[114] (Section 4.1.9.1.14)
Duplicate reflex testing without initial phlebotomy appointment, 35.8% of ptxs w/LDL[3,4]				
Duplicate reflex testing with initial phlebotomy appointment, 35.8% of ptxs w/LDL[4]				
Singlet reflex testing without initial phlebotomy appointment, 35.8% of ptxs w/LDL[3,4]				
Singlet reflex testing with initial appointment, 35.8% of ptxs w/LDL[4]				

Sensitivity analysis	Assumptions	ICER		Source (relevant sections)
		Ucar and colleagues (2017)	Arango and colleagues (2017)	
Duplicate concurrent testing without initial phlebotomy appointment[3]				
Duplicate reflex testing without initial phlebotomy appointment, 4.7% of ptxs w/LDL[3,5]				
Duplicate reflex testing with initial phlebotomy appointment, 4.7% of ptxs w/LDL[3,5]				
Singlet concurrent testing without initial phlebotomy appointment[3]				
Singlet reflex testing without initial phlebotomy appointment, 4.7% of ptxs w/LDL[3,5]				
Singlet reflex testing with initial appointment, 4.7% of ptxs w/LDL[5]				

Key: AE: adverse events; HAQ: health assessment questionnaire; ICER, incremental cost-effectiveness ratio; LDA: low disease activity; LDL: low drug level

Notes:

All costs are reported in 2017-18 prices.

1 Based on the average cost of joint replacement surgery in rheumatoid arthritis patients from the Royal Devon & Exeter NHS Foundation Trust (Appendix 8).

2 Utilities for the mixed disease population (as in the INGEBIO study) were assumed to be the same as those for people with RA

3 The cost of testing does not include the cost of an additional phlebotomy appointment which might not be required if people will be receiving regular hematological analysis as part of on-going treatement.

4 Assuming 35.8% of people have low drug level (Laine and colleagues 2016)[2]

5 Assuming 4.7% of people have low drug level (Chen and colleagues 2015)[65]

In all but one sensitivity analysis based on Ucar and colleagues (2017),[11] the intervention dominated standard care. When the impact of *flares only* was modelled (i.e. health states and AEs were not included), standard care dominated test-based treatment strategy (Table 73). In none of the sensitivity analyses based on data from Arango and colleagues (2017)[15] was there any change to the finding that standard care dominates the intervention.

4.2.1.4 Deterministic sensitivity analysis

One-way sensitivity analyses for some of the parameters used to estimate the ICERs based on data from Arango and colleagues (2017)[15] were also conducted (Table 74). Changing these parameters had no impact on the findings, standard care was estimated to dominate the intervention in all analyses.

Table 74: One-way deterministic sensitivity analyses based on data from Arango and colleagues (2017)

Parameter	Assumption	ICER	Source
Percentage of people in whom the biologic was tapered	+20% in the intervention arm and -20% in the control arm	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017)
Flare rate	-20% in the intervention arm, +20% in the control arm	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017)
Differential time in remission	+10% in the intervention arm, -10% in the control arm of the differential time in remission	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017)
Costs of managing health states	- 20%	ICER not relevant - Standard care dominates Intervention	Arango and colleagues (2017), Radner and colleagues (2014), Barbieri and colleagues (2005)

Key: ICER: incremental cost-effectiveness ratio

4.2.1.5 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis was deemed inappropriate because of a very substantial variation in clinical practice with respect to disease management in people with RA in England.

4.2.2 Etanercept and Infliximab and Promonitor

The cost-effectiveness of TNF testing in people treated with Etanercept (originator and biosimilar) and Infliximab (biosimilar) using the Promonitor test kit was explored in scenario analyses. In those analyses, it was assumed, based on clinical advice (and a lack of evidence to the contrary), that the clinical effectiveness of the different TNF inhibitors is likely to be the same, and the clinical effectiveness estimates from Ucar and colleagues (2017)[11] were adopted, with all assumptions, except acquisition and administration costs, as in Table 42. The information on the actual costs to the NHS of the TNF inhibitors was not available to the AG at the time of writing, and therefore the list prices of the biologics were assumed. The results are presented in Table 75.

Table 75: Cost-effectiveness results for the other tests and TNF inhibitors: people in remission/LDA

Treatment	Cost per year (£)	ICER Ucar and colleagues (2017)	ICER Arango and colleagues (2017)
ETN			
*Enbrel®**	9,327	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention
Erelzi	8,394	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention
IFX[1]			
Flixabi/ Renflexis (no wastage)	5164	ICER not relevant - Intervention dominates standard care	ICER not relevant - Standard care dominates Intervention

Key: ETN: etanercept; ICER: incremental cost-effectiveness ratio; IFX: infliximab

Notes:

It was assumed that blood samples would be sent for testing to UK laboratories, and the postage of £4 (per small parcel) was applied[17].

* The originator (or reference) product

[1] IFX administration cost was assumed to be 283 per injection (Section 4.1.9.1.7).

Source: Ucar and colleagues (2017)[11] and Arango and colleagues (2017)[15]

Other scenario analyses considered but not conducted due to no or low quality clinical data were: analysis of testing in the context of primary or secondary non-response; analysis for non-responders who did not adhere to treatment with biologic therapies, including switching to intravenously administered IFX.

4.3 Discussion

Despite substantial weaknesses in the clinical effectiveness evidence base (Section 2), a simple model was developed to estimate the cost-utility of ELISA test-based monitoring for people with RA taking bDMARDS.

The analyses conducted are inconclusive and suggest considerable uncertainty in the cost-effetciveness of therapeutic monitoring of TNF-alpha inhibitors in RA. Data from 2 reports of the same study produced very different conclusions on the cost-effectiveness of Promonitor testing in people receiving ADL who are in remission/LDA. The results based on the longer follow-up (Arango and colleagues 2017[15]) suggested that monitoring is more costly and produces fewer QALYs than standard care.

Of the sensitivity analyses conducted, only the assumption that the rate of flares alone changes as a consequence of monitoring, impacted on the results. This was when evidence from Ucar and colleagues (2017)[11] was used and resulted in standard care dominating the intervention.

Exploratory analyses of using Promonitor to monitor patients in remission/LDA receiving ETN or INF were undertaken, and showed the same results as that for ADL: using the longer follow-up (Arango and colleagues 2017[15]) monitoring is more costly and produces fewer QALYs than standard care.

The main effectiveness evidence in the model was from the poorly reported INGEBIO study (a non-randomised controlled trial from Spain, where <40% of participants had RA), heavily supplemented by input parameters from other studies and expert advice. The results of the economic analysis should therefore be viewed as exploratory and highly speculative. For example, although the INGEBIO study only evaluated testing using Promonitor ELISA kits, for those in remission/LDA treated with Humira® (ADL), with further assumptions these results have been used to estimate the threshold testing costs at which TDM would become cost-effective with people taking other TNF inhibitors (and taking either originator products or biosimilars.

In summary, there is much uncertainty in relation to key potential drivers of the effectiveness and cost-effectiveness of using ELISA based testing to monitoring treatment with bDMARDs in people with RA, that no firm conclusions can be drawn.

5 Discussion

5.1 Statement of principal findings

5.1.1 Clinical effectiveness

Eight studies (reported in 11 publications)[11,12,15,58-65] were included in the systematic review of the evaluation of using ELISA tests for therapeutic drug monitoring (TDM) on clinical outcomes in people with RA who had achieved remission or low disease activity [LDA], or in those who had experienced a primary non-response or a secondary non-response. Three articles[11,15,64] reported the same non-randomised controlled trial (the INGEBIO study). The remaining studies were observational studies evaluating the impact of TDM. The non-randomised controlled study[11,15,64] was judged to be at serious risk of bias. One observational study[59] had a historical control while other observational studies were single-arm studies with no comparator. For observational studies, the historical controlled study and all the single arm studies were judged to be at moderate risk of bias. However, the study design should be taken into consideration in interpreting the risk of bias assessment (non-randomised controlled study vs. observational studies).

The majority of included studies used Promonitor ELISA kits for monitoring drug levels and/or anti-drug antibody levels. Three studies in four sources (Pascual-Salcedo 2013; Paredes and colleagues (2015); Paredes and colleagues (2016**);** Lopez-Casla and colleagues (2013)[59-62] used Sanquin ELISA kits to measure drug levels and/or anti-drug antibody levels. The included studies measured drug levels and/or anti-drug antibody levels in patients who were being treated with adalimumab (ADL), etanercept (ETN) and/or infliximab (IFX). There were no studies identified for people who were being treated with certolizumab pegol (CTZ) and golimumab (GLM). No studies were identified evaluating eligible ELISA kits including IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits. Most studies enrolled rheumatoid arthritis patients who had achieved remission or low disease activities. Only one observational study (Lopez-Casla and colleagues 2013)[62] recruited people with RA who had experienced a primary non-response or a secondary non-response.

5.1.1.1 Comparative controlled evidence

Three articles[11,15,64] reported the same non-randomised controlled trial (the INGEBIO study), which focused on the population who had achieved treatment target (remission or low disease activity [LDA]). In this trial, ADL and anti-ADL antibody levels were measured using Promonitor-ADL and Promonitor-ANTI-ADL (Grifols-Progenika). Monitoring testing results were revealed to physicians in the intervention arm. Such monitoring test results were not

revealed to physicians in the control arm. This reflected standard care in Spain where treatment decisions were based on clinical judgements without the knowledge of drug levels and anti-drug antibodies of patients. This INGEBIO study recruited a mixed population of 169 including a cohort of 63 people with RA. The results of the total mixed population were reported in the review as the authors were not able to provide the results for the cohort of people with RA. The three cohorts with different conditions (rheumatoid arthritis [RA], psoriatic arthritis [PsA] and ankylosing spondylitis [AS]), may have different treatment responses to TNF-α inhibitor therapies. Therefore, there was limited generalisability of findings from this mixed population to the target RA population.

The findings from this trial (Ucar and colleagues 2017)[11] showed that, at 18-month follow-up, the rate of flares per patient-year was 0.463 for the intervention group and 0.639 for the control group, with rate difference of -0.176 (95% CI -0.379 to 0.0289). There was a non-significant reduction in risk of flare in the intervention group compared with the control group (incidence rate ratio (IRR) 0.7252, 95% CI 0.4997 to 1.0578). Median time to first flare was 145 days for participants in the intervention group and 136.5 days for participants in the control group. This trial (Ucar and colleagues, 2017[11]) further presented the results of health-related quality of life (HRQoL) outcomes. Results showed that HRQoL (EQ-5D-5L) measures were higher in the intervention group at all visits compared with the control group. However, the statistically significant results were only observed at Visit 2 (p=0.001) and Visit 3 (p=0.035). Further details of results for this outcome were not reported.

Overall, the findings from this non-randomised controlled trial (the INGEBIO study) showed that there was a non-significant reduction in risk of flare in the intervention group (where treatment decisions were made on the basis of the results of therapeutic drug monitoring) compared with the control group (i.e. standard care where treatment decisions were based on clinical judgements without the knowledge of drug levels and anti-drug antibodies of patients). HRQoL outcomes were higher in the intervention group at all visits compared with the control group, with statistically significant results being observed at two visits. However, the quality of this trial was judged to be at serious risk of bias due to potential attrition bias and baseline imbalance in disease severity between the two groups. Therefore, the results should be interpreted with caution.

5.1.1.2 Evidence from observational studies

Seven observational studies (reported in eight publications) were identified that evaluated the effect of TDM on clinical outcomes in people with RA who had achieved remission or LDA, or in those people who had experienced a primary non-response or a secondary non-response.

5.1.1.2.1 Change in disease response

Five observational studies (reported in six publications) (Chen and colleagues 2016; Inciarte-Mundo and colleagues 2016; Paredes and colleagues 2015; Paredes and colleagues 2016; Chen and colleagues 2016; Rosas and colleagues 2015)[58,60,61,63,65] assessed changes in disease response in people with RA. All studies focused on people with RA who had achieved treatment target (remission or LDA). Three uncontrolled observational studies (in four sources) (Chen and colleagues 2016; Inciarte-Mundo et al/. 2016; Paredes and colleagues 2015; Paredes and colleagues 2016)[60,61,63,65] evaluated the effect of optimisation of anti-TNF therapies (by decreasing dose or treatment frequency) by applying TDM in peple who had achieved remission or LDA. The findings showed that proportions of people who developed flares during follow-up (24 weeks to four years) ranged from 17% to 35.2%. Only one observational study (Senabre and colleagues, 2017)[12] assessed the effect of TDM in people with RA in remission receiving anti-TNFs with extended interval of administration and reported the outcome of proportions of participants who had experienced worsening of clinical activity. The finding from this study showed that 23% participants had experienced worsening of clinical activity at one-year follow-up.

The findings from two prospective uncontrolled cohort studies (Chen and colleagues 2016; Rosas and colleagues 2015)[58,65] showed that, following the anti-TNF dose tapering strategy (dose reduction), the proportions of participants who achieved persistent remission at follow-up was 87.0% and 92.0%.The study by Chen and colleagues (2016)[65] had a duration of 24-week follow-up but the study by Rosas and colleagues (2015)[58] did not report duration of follow-up. One retrospective uncontrolled cohort study (Paredes and colleagues 2016)[61] evaluated the use of a tapering strategy (dose reduction or discontinuation) of anti-TNF in people with RA with LDA or clinical remission and reported the comparative result of remission rates between pre-visit (baseline) and final visit at the duration of four-year follow-up. The results from this retrospective cohort study (Paredes and colleagues 2016)[61] showed that, in comparison with the baseline remission rate of 77%, 50% of patients maintained remission at final visit after four-year follow-up.

Overall, the evidence from these observational studies generally showed that there was a positive effect in achieving persistent remission associated with TDM for optimisation of anti-TNF therapies (by decreasing dose or treatment frequency) in people who had achieved remission or LDA. However, given that these studies were judged to be at moderate risk of bias, there were considerable uncertainties associated with the reliability of these findings.

5.1.1.2.2 Change in disease activity

Two observational studies (Pascual-Salcedo and colleagues 2013; Paredes and colleagues 2015)[59,60] evaluated the effect of TDM on change in disease activities at duration of follow-up of two to seven years, with sample sizes ranging from 43 to 54. Both studies focused on participants who had achieved remission or LDA. Pascual-Salcedo and colleagues (2013) examined two different time periods, pre- and during-TDM practice. The study showed a non-significant reduction in the mean DAS28 score following the implementation of TDM at seven-year follow-up (pre-TDM: mean 2.51 [SD 0.85] vs.during-TDM: 2.31, [SD 0.52]; p=0.061).

Another observational study by Paredes (2015)[60] assessed the outcome measure of DAS28 score in patients receiving TNF-inhibitor therapies at pre-visit (baseline) and post-visit with 2-year follow-up. The results showed that therapeutic drug monitoring for optimisation of anti-TNF therapies was associated with a non-significant reduction in mean DAS28 score at post-visit after 2-year follow-up compared with pre-visit in patients receiving infliximab therapies, but non-signficant increases in mean DAS28 scores at 2-year follow-up were observed in those patients receiving adalimumab and etanercept.

Overall, the finding from the historically controlled study (Pascual-Salcedo 2013)[59] showed that therapeutic drug monitoring was associated with a non-significant reduction in mean DAS28 scores at seven-year follow-up compared with the historical control period. However, mixed results were found in a retrospective uncontrolled cohort study by Paredes (2015).(52) Given the inconsistency of results, there was uncertainty on the impact of TDM on disease activity. It should be noted that the quality of data was judged to be at moderate risk of bias, which compromises the reliability of the findings.

5.1.1.2.3 Change in direction and magnitude of therapeutic dose

Three observational studies (Pascual-Salcedo 2013; Rosas 2015; Paredes 2016)[58,59,61] evaluated the outcome of changes in direction and magnitude of therapeutic dose in people with RA who had achieved remission or LDA. The sample size of included studies ranged from 43 to 52.

The findings from the study by Pascual-Salcedo (2013)[59] demonstrated that, compared with the historical control period without TDM, there were statistically significant reductions in weekly mean dose per patient by each anti-TNF (AFX, ADL, ETN) during the 2nd period where TDM was introduced. The findings from this study further showed that, compared with the historical control, there were statistically significant increases in the mean interval of administration for each anti-TNF during the 2nd period when TDM was implemented.

Only one prospective observational study by Rosas and colleagues (2015)[58] assessed the impact of TDM on the total dose of anti-TNFs avoided. The results demonstrated that the total number of doses avoided was 548 for ETN therapies and 260 for ADL therapies, respectively. However, this study did not report the duration of follow-up. Another retrospective observational study by Paredes and colleagues (2016)[61] assessed the mean drug levels between the pre-visit (baseline) and post-visit (follow-up) for each anti-TNF (adalimumab, etanercept and infliximab) at the duration of four-year follow-up. The results showed that, compared with baseline measurements, there were statistically significant reductions in mean drug levels at post-visit for each anti-TNF evaluated.

Overall, the limited data from three observational studies showed that TDM for optimising anti-TNF therapies was associated with reductions in therapeutic dose of anti-TNFs in people with RA who had achieved remission or LDA. This would be expected to lead to cost saving associated with TDM. However, the reliability of the findings may be compromised by the poor quality of the studies.

5.1.1.2.4 Discontinuation of ineffective therapy

There were limited data identified for the assessment of an impact of therapeutic drug monitoring on treatment decision making and management. Among included studies, only one prospective observational study by Lopez-Casla and colleagues (2013)[62] assessed the impact of therapeutic drug monitoring on treatment decision making and reported the outcome of treatment discontinuation. This study assessed whether therapeutic drug monitoring for optimising anti-TNF therapies (e.g. increasing the dose of IFX) was an effective therapeutic strategy in 36 people with RA who developed clinical inefficacy (i.e. participants who had experienced a primary non-response or a secondary non-response). The study by Lopez-Casla and colleagues (2013)[62] reported that 76.5% of participants discontinued their anti-TNF therapies (IFX) in the one-year follow-up period due to ineffectiveness of anti-TNF therapies. The study concluded that increasing the dose of IFX did not seem to be an effective option in people with RA who had developed clinical inefficacy.

5.1.2 Cost effectiveness

The analyses conducted are inconclusive and suggest considerable uncertainty in the cost-effetciveness of therapeutic monitoring of TNF-alpha inhibitors in RA. Data from 2 reports of the same study produced very different conclusions on the cost-effectiveness of Promonitor testing in people receiving ADL who are in remission/LDA. The results based on the longer

follow-up (Arango and colleagues 2017[15]) suggested that monitoring is more costly and produces fewer QALYs than standard care.

Of the sensitivity analyses conducted, only one assumption impacted on the results: that the rate of flares alone changes as a consequence of monitoring. This was when evidence from Ucar and colleagues (2017)[11] was used and resulted in standard care dominating the intervention.

Exploratory analyses of using Promonitor to monitor patients in remission/LDA receiving ETN or INF were undertaken, and showed the same results as that for ADL: using the longer follow-up (Arango and colleagues 2017[15]) monitoring is more costly and produces fewer QALYs than standard care.

5.2 Strengths and limitations of the assessment

5.2.1 Clinical effectiveness

Extensive literature searches were conducted with an attempt to maximize the retrieval of potentially relevant studies for the systematic review of clinical effectiveness. These included electronic searches of a variety of bibliographic databases as well as screening of clinical trial registers and conference proceedings to identify unpublished studies. The search strategy did not restrict by study design. The review process followed recommended methods to minimise the potential for error and/or bias. The quality of included studies was assessed and accounted for when interpreting the review results. Appropriate synthesis methods were employed by taking into account the heterogeneity of study characteristics.

In terms of limitations, only studies in English were included, therefore some potentially relevant non-English language studies may have been missed. There was scarce evidence relating to clinical effectiveness of TDM on clinical outcomes in people RA who had experienced a primary non-response or a secondary non-response. No studies were identified assessing ELISA kits including IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits. There was considerable clinical heterogeneity associated with interventions, outcomes and length of follow-up between included studies. We were unable to investigate publication bias, because quantitative synthesis was not possible in this systematic review due to considerable clinical heterogeneity.

5.2.2 Cost effectiveness

A systematic review of published economic evaluations of using ELISA tests relative to the alternatives and standard care was undertaken to help inform the type and structure of the

decision model. The results of this review indicate limited evidence on the cost-effectiveness of TDM in people with RA. Despite a comprehensive search of the literature, only five studies have been identified. Two (out of five) TNF testing kits from the NICE scope (Promonitor and Sanquin) and three (out of five) TNF inhibitors (ADL, ETN, IFX) have been assessed in the selected studies. The systematic review was also limited by reporting as two (out of five) selected studies were reported in the abstract form.

Only in the INGEBIO study, selected in the clinical-effectiveness systematic review, was a test-based treatment compared against usual care. In this study, however, physicians were not obliged to follow any test-based treatment algorithm but could use testing to alter doses, based on their judgement, in patients from the intervention arm. Moreover, the study was reported in the abstract form only, and the reported outcomes were not directly relevant to the NHS since the study was conducted in Spain. Therefore, an additional systematic literature review to identify RCTs evaluating *any tests* used to monitor anti-TNF-α treatment of people with RA was conducted to support the economic assessment. However, no relevant sources were identified.

Due to the limited evidence available on clinical effectiveness of TNF monitoring in people with RA, the multifactorial nature of decisions to adjust treatments in people with RA,[50] and the recent changes in the biologics market, which contributed to the uncertainty in the prices of biologics and their uptake within the UK, a simplified modelling approach, a threshold analysis, was chosen to address the decision problem. In this analysis, the cost of measuring drug concentrations and anti-drug antibody levels at which addition of TNF testing to usual practice is likely to have zero net monetary benefit (NMB), was estimated in people with RA treated with ADL for a range of annual acquisition costs. The estimates obtained under the cost-effectiveness thresholds of £20,000 per QALY gained and £30,000 per QALY gained were compared against those derived from literature and provided to the AG by our external advisors.

The most important limitations of the economic analysis undertaken in this study are described below:

- The major challenge in this assessment was limited evidence on clinical effectiveness, health-related quality of life (HRQoL) and costs associated with test-based treatment strategies. Due to the paucity of data, not all test kits and TNF inhibitors, and not all populations specified in the NICE scope were considered in the economic analysis. In particular, the evidence related to primary non-responders and secondary non-responders, identified in the clinical-effectiveness systematic review,

did not directly compare the intervention under evaluation to the alternative, and this evidence was not sufficient for conducting an economic analysis relevant to non-responders. Moreover, no economic evaluations relevant to IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits have been conducted. Furthermore, separate economic assessment of testing for free and total anti-drug antibodies was not possible since it was unclear whether free anti-drug antibody assays or total anti-drug antibody assays were performed in the selected studies.

- Several test–based treatment algorithms have been proposed and used, by physicians in the UK; e.g. Exeter biologic clinic recommendations for biologic dose reduction (Appendix 5) and recommendations by NHS Greater Glasgow and Clyde on biologic drug monitoring (Appendix 6). However, to our knowledge, there is no unified treatment algorithm based on TNF testing. Importantly, in the INGEBIO study (conducted in Spain), clinicians were not expected to follow any test-based strategy when making treatment decisions based on test results and clinical judgement. Therefore, it is unclear whether and to what extent the economic results based on this study are relevant to clinical practice in England.
- To our knowledge, there is no unified recommendation on managing flares in the UK people with RA. To address this limitation, several sensitivity analyses informed by literature and based on clinical expert advice were carried out. It is not clear, however, which of those analyses is most relevant to the NHS.
- The time horizon of the threshold analysis undertaken in this study was defined by the observational period in the INGEBIO trial, which was conducted for 18 months. Cost and health outcomes were not extrapolated into the future, since, due to the lack of long-term clinical studies, external validation of extrapolated outcomes would not be feasible. Furthermore, given multifactorial nature of treatment decisions in people with RA, long-term extrapolation of cost and health outcomes would be prone to even greater uncertainties, which would not be possible to quantify given substantial limitations in the evidence base.
- Due to limited reporting, it is not clear to what extent selection bias in the INGEBIO study (which was a non-randomised trial) could have influenced the results of the economic analysis.
- In this study, as in many other economic evaluations in RA, health state utility values were estimated from HAQ scores using published regression functions. These functions have demonstrated a relatively strong correlation between the HAQ and several HRQoL instruments. The AG adopted this approach since the evidence on

HRQoL from the INGEBO study was limited. We recognise, however, that the HAQ is a functional measure, and does not capture the full impact of RA on quality of life.

- Utility values estimated from HRQoL data for people with RA were applied based on clinical outputs from the INGEBIO study which had a *mixed* population of people with RA, PsA, and AS. Since people with RA are usually older and more likely to be female when compared with people with PsA or AS, the utility values for people with RA, used in the economic analysis, are likely to be lower than those for the mixed population (since men tend to value health states higher than women, and the same applies to younger versus older people).[18] This may have overestimated the incremental cost-effectiveness ratios (ICERs).
- Since the rates of AE were not reported in the INGEBIO study, the impact of AEs was modelled using evidence from another study, which is a limitation of this analysis. However, based on clinical advice and published literature on adverse events in people with RA treated with TNF inhibitors, those AEs which carry a significant cost and disutility burden are relatively rare.
- Finally, limited evidence on utilities, based on EQ-5D scores, directly relevant to people with flares, people experiencing serious adverse events as well as people with remission/LDA or active disease health status in the UK settings was identified in this study. Therefore, utilities were derived from HAQ scores which were estimated in studies conducted in people with RA in non-UK settings. It should be noted however that utilities were estimated by mapping to EQ-5D outcomes from UK tariffs.

5.3 Uncertainties

5.3.1 Clinical effectiveness

In this assessment we identified limited data that evaluated clinical effectiveness of using ELISA tests for monitoring response to TNF-α inhibitors in people with RA who had achieved remission or LDA, or in those people who had experienced a primary non-response or a secondary non-response. There were scarce data identified for people who had experienced a primary non-response or a secondary non-response. In particular, we did not identify any RCTs evaluating patient-related outcomes and disease activities associated with using ELISA tests for TDM in the target populations.

The non-randomised controlled study[11,15,64] was judged to be at serious risk of bias. For observational studies, the historically controlled study and all the single arm studies were judged to be at moderate risk of bias. In the non-randomised controlled trial (the INGEBIO study), there was baseline imbalance in disease severity between the intervention and control groups. Furthermore, there was a lack of adjusting for this variable in the analysis of

clinical outcomes. There were high attrition rates for some outcomes, which could lead to attrition bias. The historically controlled study by Pascual-Salcedo and colleagues (2013)[59] was associated with non-contemporaneous control bias due to the use of a historical control. Most observational studies had a small sample size without a control group. Given the poor quality of included studies, the potential role of ELISA testing in terms of its clinical impact on monitoring response to TNF-α inhibitors in the target populations remains unclear.

5.3.2 Cost effectiveness

Since there is neither gold standards nor guidelines available to monitor the TNF inhibitors considered in this assessment, economic analyses of test-based treatment strategies with biologics represent a substantial challenge.

Due to data limitations and the lack of clarity with regard to test-based treatment strategies, the AG deliberately refrained from data-intensive modelling approaches, which would be impossible to implement without making strong assumptions not supported by evidence.

The studies identified in the clinical-effectiveness systematic review and used to inform the model structure and parameters are limited by study design (e.g. none of the studies were randomised, seven out of eight studies were observational). Furthermore, those studies are characterised by relatively small sample sizes. Only in one study (INGEBIO), treatment of RA patients based on the results of TNF testing was compared against usual care, and this was in a *mixed disease* population with only 37% of RA patients. The overall majority of the selected studies, including the INGEBIO study, were reported as abstracts only. Of note, the studies were sponsored by pharmaceutical companies.

The AG is aware of several test-based treatment algorithms used by physicians in England. However, in the only study comparing test versus no-test treatment strategies, the INGEBIO study (which was utilised in our economic analysis), physicians were not required to follow any therapeutic algorithm based on TDM results but could use tests to alter doses based on their clinical judgement. It is unclear, however, whether there are variations in clinical practice between England and Spain, which could have impacted the results presented here.

Only three studies included in the clinical-effectiveness systematic review reported the definitions of flare, and those definitions were consistent between the studies (DAS28 > 3.2). However, the AG is aware of several RA flare criteria, which have been used in clinical research.[9,67] Our clinical advisors confirmed that such a variation also exists in clinical practice. To address this uncertainty, the effect of the variation in the definition of flare was

explored in a number of sensitivity analyses by altering assumptions on the duration of flare and the effect of flare on health-related quality of life.

In the INGEBIO study, the rates of adverse events in the intervention and control arms were not reported. Therefore, the impact of AEs on costs and QALYs was investigated assuming AE rates from another study selected in the systematic review of clinical effectiveness.

Since the INGEBIO study was carried out in Spain, and the reported outcomes (average acquisition costs per patient-year, and QALYs accrued over the duration of the study) were not directly relevant to the NHS, some important assumptions had to be made in the analyses conducted by the AG. In particular, it was assumed that clinical practice in England with respect to treatment decisions in people with RA on biologics is similar to that in Spain.

Finally, since the actual costs to the NHS of adalimumab (Humira®), its biosimilars and other TNF treatments were not known to the AG at the time of writing, the effect of variation in the annual acquisition costs of the biologics within the range of £1,000-£9,200 per patient was examined in the threshold analysis. However, given the fact that: (1) the actual costs of the originator products and their biosimilars vary considerably across England, (2) there is also a variation in the uptake of biosimilars across the UK, and (3) the proportion of people treated with biosimilars is likely to increase in the near future due to very recent changes in the biologics market, it is not clear which estimates obtained in our economic analyses are most relevant to the NHS.

5.4 Generalisability of the findings

5.4.1 Clinical effectiveness

Given that most studies were conducted in Spain, the generalisability of their findings to the UK settings remains uncertain due to variations in clinical practice and health policies between different countries. Furthermore, the findings from the non-randomised controlled trial (the INGEBIO study) and the results of changes in therapeutic dose from the study by Pascual-Salcedo and colleagues (2013)[59] were presented for a mixed population. Therefore, there was limited generalisability of findings from the mixed population (including RA, PsA, and/or AS) to the target RA population.

5.4.2 Cost effectiveness

Outcomes from the INGEBIO study were utilised in the economic analysis for patients in remission/LDA. It was a pragmatic trial, and therefore it is likely that the results could be generalisable to routine practice settings. However, the generalisability to the UK settings of

the findings in the INGEBIO study and the economic results reported in this assessment remain uncertain due to likely variations in clinical practice between Spain and England.

Since findings from the mixed population considered in the INGEBIO study might not be generalisable to the RA population, and the quality of this trial was judged to be at serious risk of bias, the economic results presented here should be considered with caution.

Due to the paucity of data, not all test kits and TNF inhibitors from the NICE scope could be modelled using reported clinical outcomes considered in this study, and it is not clear whether and to what extent the economic estimates obtained for patients treated with adalimumab are applicable to people treated with the other anti-TNF treatments.

Moreover, data limitations did not allow the assessment of the long-term economic impact of TNF testing since TDM in people with RA is relatively new, and therefore there is no data relevant to the long-term outcomes of test-based treatment strategies. Given the dynamic nature of RA treatment and limited data, it is not known whether the reported clinical effects and associated incremental costs of test-based treatment decisions would persist beyond this time.

According to NHS England,[91] some originator manufacturers have offered discounts, further enhancing the competitiveness of the market and potential for cost saving for the NHS. Therefore, the list prices of TNF inhibitors assumed in the analyses reporting ICERs (Table 71, Table 73, Table 74 and Table 75) might not adequately reflect the actual costs of the biologics to the NHS.

6 Conclusions

6.1 Implications for service provision

The findings from this assessment demonstrate very limited evidence on the effect of TDM based on ELISA tests for optimising anti-TNF therapies in people with RA, either in those who had achieved remission or LDA, or in those who had experienced a primary non-response or a secondary non-response.

In relation to clinical effectiveness, limited data were identified evaluating TDM in the target populations. One non-randomised trial compared TDM with standard care (the INGEBIO study) had serious limitations in relation to the NICE scope: only one-third of the participants had RA, many of the analyses were not by intention-to-treat, follow-up was for only 18 months, there was no explicit algorithm for guiding clinicians in how the results of testing should change treatment (e.g. tapering), and the study was only reported in three abstracts. In addition, seven observational studies (reported in eight publications) were also identified but were of minimal value in informing whether ELISA test-based monitoring is clinically effective or not.

Despite these substantial weaknesses in the clinical effectiveness evidence base, a simple model was developed to estimate the cost-utility of ELISA test-based monitoring for people with RA taking bDMARDS. The main effectiveness evidence in the model was also from the poorly reported INGEBIO study, heavily supplemented by input parameters from other studies and expert advice. The results of the economic analysis should therefore be viewed as exploratory and highly speculative. For example, although the INGEBIO study only evaluated testing using Promonitor ELISA kits, for those in remission/LDA treated with Humira® (ADL), with further assumptions these results have been used to estimate the threshold testing costs at which TDM would become cost-effective with people taking other TNF inhibitors (and taking either originator products or biosimilars.

In summary, there is limited valid and applicable research evidence, and much uncertainty in relation to key potential drivers of the effectiveness and cost-effectiveness of using ELISA based testing to monitoring treatment with bDMARDs in people with RA, that no firm conclusions regarding the implications for service provision can be drawn

6.2 Suggested research priorities

One ongoing Norwegian multicentre randomised controlled trial (the NOR-DRUM Study)[66] that evaluates the effect of TDM in people with RA in remission compared with standard

care. This ongoing trial will provide further useful data on the impact of TDM in the target population.

Further controlled trials with a large sample size (especially randomised controlled trials [RCTs]) are required to assess the impact of using Promonitor ELISA tests for monitoring anti-TNF therapies in people with RA who had achieved remission or LDA.

No studies were identified evaluating other eligible ELISA kits including IDKmonitor ELISA kits, LISA-TRACKER ELISA kits, RIDASCREEN ELISA kits and MabTrack ELISA kits. Therefore, future large RCTs are required to assess the impact of using those ELISA tests for monitoring anti-TNF therapies in people with RA who had achieved remission or LDA. More robust evidence is also needed to evaluate the impact of using Sanquin tests for monitoring anti-TNF therapies in this population.

There were no studies identified for people with RA treated with CTZ or GLM. Future RCTs are required to assess the clinical effectiveness of using ELISA tests for monitoring such anti-TNF therapies in the target populations.

There were scarce data identified for the population of people with RA who had experienced a primary non-response or a secondary non-response. Future RCTs are warranted to assess the clinical effectiveness of using ELISA tests for monitoring anti-TNF therapies in those who had developed clinical inefficacy.

Limited evidence on healthcare resource use and utilities, based on EQ-5D scores, directly relevant to the population considered in this assessment was identified in this study. This warrants further research on medium/long term cost and health outcomes in people with RA treated with TNF inhibitors.

7 References

1. Krieckaert CLM, Nair SC, Nurmohamed MT, van Dongen CJJ, Lems WF, Lafeber FPJG, et al. Personalised treatment using serum drug levels of adalimumab in patients with rheumatoid arthritis: an evaluation of costs and effects. Annals of the Rheumatic Diseases. 2015;74(2):361-8.

2. Laine J, Jokiranta TS, Eklund KK, Väkeväinen M, Puolakka K. Cost-effectiveness of routine measuring of serum drug concentrations and anti-drug antibodies in treatment of rheumatoid arthritis patients with TNF-α blockers. Biologics. 2016;10:67-73.

3. Bingham CO, Pohl C, Woodworth TG, Hewlett SE, May JE, Rahman MU. Developing a standardized definition for disease 'Flare' in rheumatoid arthritis (OMERACT 9 special interest group). J Rheumatol. 2009;36.

4. Maravic M BC, Daurès JP et al. Practices for managing a flare of long-standing rheumatoid arthritis: Survey among French rheumatologists Clinical and Experimental Rheumatology. 2005;23(36):36-42.

5. Barbieri M, Wong JB, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. PharmacoEconomics. 2005;23(6):607-18.

6. Radner H, Smolen JS, Aletaha D. Remission in rheumatoid arthritis: benefit over low disease activity in patient-reported outcomes and costs. Arthritis Research & Therapy. 2014;16(1).

7. NHS. Reference costs 2017 [Available from: https://improvement.nhs.uk/resources/reference-costs/.

8. National Insitute of Health and Care Excellence (NICE). Adalimumab, etanercept, inflfliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed; nice.org.uk/guidance/ta375. 2016.

9. Markusse IM, Dirven L, Gerards AH, van Groenendael JHLM, Ronday HK, Kerstens PJSM, et al. Disease flares in rheumatoid arthritis are associated with joint damage progression and disability: 10-year results from the BeSt study. Arthritis Research & Therapy. 2015;17(1):232.

10. Oppong R, Kaambwa B, Nuttall J, Hood K, Smith RD, Coast J. The impact of using different tariffs to value EQ-5D health state descriptions: an example from a study of acute cough/lower respiratory tract infections in seven countries. The European Journal of Health Economics. 2013;14(2):197-209.

11. Ucar E, Gorostiza I, Gomez C, Perez C, Dios J, Alvarez B, et al. Prospective, intervention, multicenter study of utility of biologic drug monitoring with respect to the efficacy and cost of adalimumab tapering in patients with rheumatic diseases: preliminary results of ingebio study. Annals of the rheumatic diseases Conference: annual european congress of rheumatology, EULAR 2017 Spain [Internet]. 2017; 76(Supplement 2):[826 p.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/706/CN-01467706/frame.html.

12. Senabre Gallego JM, Rosas Gomez De Salazar J, Marco Mingot M, Naranjo A, Llinares-Tello F, Pons A, et al. Clinical activity, ultrasound assessment and drug monitoring

in rheumatoid arthritis patients receiving anti-TNF-alpha therapy with extended interval of administration. Annals of the Rheumatic Diseases. 2017;76 (Supplement 2):1016.

13. Singh JA, Cameron C, Noorbaloochi S, Cullis T, Tucker M, Christensen R, et al. The risk of serious infection with biologics in treating patients with rheumatoid arthritis: A Systematic Review and Meta-analysis. Lancet. 2015;386(9990):258-65.

14. Malottki K, Bartos P, Tsourapas A, Uthman OA, 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 Technology Assessment. 2011;15(14):1-278.

15. Arango CG, Vivar MLG, Angulo EU, Gorostiza I, Perez CE, De Dios JR, et al. Prospective, intervention, multicenter, non-inferiority study of utility of therapeutic drug monitoring with respect to the efficacy and cost of adalimumab tapering in patients with rheumatic diseases. Arthritis and Rheumatology Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP. 2017;69(Supplement 10).

16. Bykerk VP, Shadick N, Frits M, Bingham CO, Jeffery I, Iannaccone C. Flares in rheumatoid arthritis: frequency and management. A report from the BRASS registry. J Rheumatol. 2014;41.

17. Mail R. Get a price 2018 [Available from: https://www.royalmail.com/price-finder/.

18. Ara R, Brazier JE. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. Value Health. 2010;13(5):509-18.

19. National Audit Office. Services for people with rheumatoid arthritis. REPORT BY THE COMPTROLLER AND AUDITOR GENERAL | HC 823 Session 2008-2009 | 15 July 2009. 2009.

20. Dickson JL, P; Wise, E. Making a Diagnosis of Rheumatoid Arthritis 2014 [

21. Picerno V, Ferro F, Adinolfi A, Valentini E, Tani C, Alunno A. One year in review: the pathogenesis of rheumatoid arthritis. Clin Exp Rheumatol. 2015;33(4):551-8.

22. Scott IC, Steer S, Lewis CM, Cope AP. Precipitating and perpetuating factors of rheumatoid arthritis immunopathology: linking the triad of genetic predisposition, environmental risk factors and autoimmunity to disease pathogenesis. Best Pract Res Clin Rheumatol. 2011;25(4):447-68.

23. Symmons D. What is the Cause of Rheumatoid Arthritis? Non-Genetic Factors 2013 [

24. Karlson EW, Mandl LA, Hankinson SE, Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' Health Study. Arthritis Rheum. 2004;50(11):3458-67.

25. Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A, et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. Ann Rheum Dis. 2010;69(1):70-81.

26. Society NRA. The genetics of rheumatoid arthritis 2018 [Available from: www.nras.org.uk/the-genetics-of-rheumatoid-arthritis.

27. Philippou E, Nikiphorou E. Are we really what we eat? Nutrition and its role in the onset of rheumatoid arthritis. Autoimmun Rev. 2018;17(11):1074-7.

28. James D, Young A, Kulinskaya E, Knight E, Thompson W, Ollier W, et al. Orthopaedic intervention in early rheumatoid arthritis. Occurrence and predictive factors in an inception cohort of 1064 patients followed for 5 years. Rheumatology (Oxford). 2004;43(3):369-76.

29. Raheel S, Matteson EL, Crowson CS, Myasoedova E. Improved flare and remission pattern in rheumatoid arthritis over recent decades: a population-based study. Rheumatology (Oxford). 2017;56(12):2154-61.

30. National Institute for Health and Care Excellence (NICE). Rheumatoid arthritis in adults: management; nice.org.uk/guidance/ng100. National Institute for Health and Care Excellence; 2018.

31. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. 2010;69(9):1580-8.

32. 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 (Oxford). 2002;41(7):793-800.

33. Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. Br J Rheumatol. 1994;33(8):735-9.

34. Kearsley-Fleet L, Davies R, De Cock D, Watson KD, Lunt M, Buch MH, et al. Biologic refractory disease in rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. Ann Rheum Dis. 2018;77(10):1405-12.

35. Gavan S. An economic evaluation of a biomarker test to stratify treatment for rheumatoid arthritis: The University of Manchester; 2017.

36. Abhishek A, Nakafero G, Kuo CF, Mallen C, Zhang W, Grainge MJ, et al. Rheumatoid arthritis and excess mortality: down but not out. A primary care cohort study using data from Clinical Practice Research Datalink. Rheumatology (Oxford). 2018;57(6):977-81.

37. Hyrich K, Symmons D, Watson K, Silman A, Consortium BCC, British Society for Rheumatology Biologics R. Baseline comorbidity levels in biologic and standard DMARD treated patients with rheumatoid arthritis: results from a national patient register. Ann Rheum Dis. 2006;65(7):895-8.

38. National Rheumatoid Arthritis Society (NRAS). The economic burden of rheumatoid arthritis. 2010.

39. Brown EE, S; Haigh, R; Mascarenhas, R. Adherence to Biologic Therapies in Patients with Inflammatory Arthritis. Rheumatology (Oxford). 2015;54(Suppl 1):i71.

40. National Insitute of Health and Care Excellence (NICE). https://pathways.nice.org.uk/pathways/rheumatoid-arthritis#path=view%3A/pathways/rheumatoid-arthritis/drug-treatment-for-rheumatoid-arthritis.xml&%2338%3Bcontent=view-index&content=view-index [

41. Cantini F, Benucci M. Focus on biosimilar etanercept - bioequivalence and interchangeability. Biologics. 2018;12:87-95.

42. Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). Clin Exp Rheumatol. 2005;23(5 Suppl 39):S14-8.

43. 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(1):44-8.

44. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol. 2005;23(5 Suppl 39):S100-8.

45. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum. 1995;38(6):727-35.

46. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum. 1996;39(1):34-40.

47. Simpson E, Hock E, Stevenson M, Wong R, Dracup N, Wailoo A, et al. What is the added value of ultrasound joint examination for monitoring synovitis in rheumatoid arthritis and can it be used to guide treatment decisions? A systematic review and cost-effectiveness analysis. Health Technol Assess. 2018;22(20):1-258.

48. Kiely P. The DAS28 score 2017 [

49. Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol. 2005;23(5 Suppl 39):S93-9.

50. 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(35):1-610.

51. Vogelzang EH, Hebing RCF, Nurmohamed MT, van Kuijk AWR, Kruijff JWF, l'Ami MJ, et al. Adherence to etanercept therapy in rheumatoid arthritis patients during 3 years of follow-up. PLoS One. 2018;13(10):e0205125.

52. National Insitute of Health and Care Excellence (NICE). Adalimumab, etanercept, infliximab,rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor; nice.org.uk/guidance/ta195. 2010.

53. Jani M, Gavan S, Chinoy H, Dixon WG, Harrison B, Moran A, et al. A microcosting study of immunogenicity and tumour necrosis factor alpha inhibitor drug level tests for therapeutic drug monitoring in clinical practice. Rheumatology (Oxford). 2016;55(12):2131-7.

54. The Cochrane Collaboration. Cochrane handbook for systematic reviews of interventions: John Wiley & Sons, Ltd; 2008.

55. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine. 2009;6(7):e1000097.

56. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ (Clinical research ed). 2011;343:d5928.

57. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ (Clinical research ed). 2016;355:i4919.

58. Rosas J, Llinares-Tello F, Miguel Senabre J, Santos-Soler G, Salas-Heredia E, Barber X, et al. Economic impact of decreasing adalimumab and etanercept doses and drug monitoring in patients with rheumatoid arthritis in clinical remission: Preliminary study from a local biologics unit. Arthritis and Rheumatology Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP. 2015;67(SUPPL. 10).

59. Pascual-Salcedo D, Plasencia C, Gonzalez Del Valle L, Lopez Casla T, Arribas F, Villalba A, et al. Therapeutic drug monitoring (TDM) in rheumatic day clinic enables to reduce pharmaceutical cost maintaining clinical efficacy. Annals of the Rheumatic Diseases Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR. 2013;72(SUPPL. 3).

60. Paredes B, Plasencia C, Pascual-Salcedo D, Monjo I, Pieren A, Moral E, et al. Influence of optimization of biological therapies on immunogenicity in a cohort of rheumatoid arthritis with low disease activity. Annals of the Rheumatic Diseases. 2015;2):458.

61. Paredes B, Plasencia C, Pascual-Salcedo D, Monjo I, Pieren A, Moral E, et al. Influence of tapering biological therapies in immunogenicity in a cohort of rheumatoid arthritis with low disease activity. Annals of the Rheumatic Diseases. 2016;75 (Supplement 2):1003.

62. Lopez-Casla MT, Pascual-Salcedo D, Plasencia C, Alcozer P, Garcia-Carazo S, Bonilla G, et al. The infliximab dose increase is not correlated with clinical improvement in RA patients. Annals of the Rheumatic Diseases Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR. 2013;72(SUPPL. 3).

63. Inciarte-Mundo J, Hernandez M, Ruiz-Esquide V, Ramirez J, Cuervo A, Cabrera-Villalba S, et al. Prediction of flare in rheumatoid arthritis and psoriatic arthritis patients with low disease activity receiving TNF inhibitors: Role of calprotectin and drug trough serum levels. A one-year prospective cohort study. Annals of the Rheumatic Diseases. 2016;75 (Supplement 2):968-9.

64. Gorostiza I, Angulo EU, Arango CG, Perez CE, De Dios JR, Alvarez B, et al. Prospective, intervention, multicenter study of utility of biologic drug monitoring with respect to the efficacy and cost of adalimumab tapering in patients with rheumatic diseases (34-week descriptive data). Arthritis and Rheumatology. 2016;68 (Supplement 10):835-6.

65. Chen DY, Chen YM, Hsieh TY, Hung WT, Hsieh CW, Chen HH, et al. Drug trough levels predict therapeutic responses to dose reduction of adalimumab for rheumatoid arthritis patients during 24 weeks of follow-up. Rheumatology (Oxford). 2016;55(1):143-8.

66. Haavardsholm EA. The Norwegian Drug Monitoring Study (NOR-DRUM). 2018.

67. van der Maas A, Lie E, Christensen R, Choy E, de Man YA, van Riel P, et al. Construct and criterion validity of several proposed DAS28-based rheumatoid arthritis flare criteria: an OMERACT cohort validation study. Ann Rheum Dis. 2013;72(11):1800-5.

68. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. International journal of technology assessment in health care. 2005;21(2):240-5.

69. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24(4):355-71.

70. Krieckaert C, Nair SC, Nurmohamed MT, Van Dongen CJ, Lems WF, Lafeber FP, et al. Evaluating the cost-effectiveness of personalized treatment with adalimumab using serum drug level and anti-adalimumab antibodies in rheumatoid arthritis patients. Annals of the Rheumatic Disease Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR. 2012;71(SUPPL. 3).

71. Krieckaert CLM, Nair SC, Nurmohamed MT, van Dongen CJJ, Lems WF, Lafeber FPJG, et al. Evaluating the Cost-Effectiveness of Personalized Treatment with Adalimumab Using Serum Drug Levels in Rheumatoid Arthritis Patients. Arthritis Rheum-Us. 2012;64(10):S779-S.

72. Krieckaert C, Nair SC, Nurmohamed MT, van Dongen CJ, Lems WF, Lafeber FP, et al. Evaluating the Cost-Effectiveness of Personalized Treatment with Adalimumab Using Serum Drug Level and Anti-Adalimumab Antibodies in Rheumatoid Arthritis Patients. Annals of the Rheumatic Diseases. 2013;71:104-.

73. Krieckaert CLM, Nair SC, Nurmohamed MT, van Dongen CJJ, Lems WF, Lafeber FPJG, et al. Personalised treatment using serum drug levels of adalimumab in patients with rheumatoid arthritis: An evaluation of costs and effects. Annals of the Rheumatic Diseases. 2013;21.

74. Bartelds GM, Krieckaert CLM, Nurmohamed MT, van Schouwenburg PA, Lems WF, Twisk JWR, et al. Development of antidrug antibodies against adalimumab and association with disease activity and treatment failure during long-term follow-up. Jama-J Am Med Assoc. 2011;305(14):1460-8.

75. Vincent FB, Morand EF, Murphy K, Mackay F, Mariette X, Marcelli C. Antidrug antibodies (ADAb) to tumour necrosis factor (TNF)-specific neutralising agents in chronic inflammatory diseases: a real issue, a clinical perspective. Annals of the Rheumatic Diseases. 2013;72(2):165-78.

76. Smolen JS, Jones H, Mahgoub E, Pedersen R, Marshall L. Association between Flare and Radiographic Progression in Patients with Rheumatoid Arthritis. Annals of the Rheumatic Diseases. 2017;76:832-.

77. Dennison EM, Packham J, Hyrich K. The BSRBR-RA at 15 years. Rheumatology (Oxford). 2016;55(12):2093-5.

78. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. NICE Process and Methods Guides. 2013.

79. Lahiri M, Dixon WG. Risk of infection with biologic antirheumatic therapies in patients with rheumatoid arthritis (vol 29, pg 290, 2015). Best Pract Res Cl Rh. 2015;29(6):826-.

80. Bruce ES, Kearsley-Fleet L, Watson KD, Symmons DP, Hyrich KL. Risk of Pneumocystis jirovecii pneumonia in patients with rheumatoid arthritis treated with inhibitors of tumour necrosis factor α: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. Rheumatology (Oxford). 2016;55(7):1336-7.

81. Dixon WG, Hyrich K, Watson K, 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(3):522-8.

82. Burmester GR, Landewé R, Genovese MC, Friedman AW, Pfeifer ND, Varothai NA, et al. Adalimumab long-term safety: infections, vaccination response and pregnancy outcomes in patients with rheumatoid arthritis. Ann Rheum Dis. 2016;76(2).

83. Singh JA, Wells GA, Christensen R, Tanjong Ghogomu E, Maxwell L, Macdonald JK, et al. Adverse effects of biologics: a network meta-analysis and Cochrane overview. Cochrane Database Syst Rev. 2011(2):CD008794.

84. Jani M, Dixon WG, Lunt M, De Cock D, Isaacs JD, Morgan AW, et al. The association of biologic drug-levels with infection risk: results from the british society for rheumatology biologics register for rheumatoid arthritis. Ann Rheum Dis. 2018;77(Suppl 2).

85. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. NICE DSU Technical Support Document 4: Inconsistency in networks of evidence based on randomised controlled trials. Sheffield: Decision Support Unit, School of Health and Related Research (ScHARR), University of Sheffield; 2011 (last updated April 2014).

86. Joint Formulary Committee. British National Formulary (online) London: BMJ Group and Pharmaceutical Press <http://www.medicinescomplete.com> [Accessed on [03/12/18]] [

87. The Campbell and Cochrane Economics Methods Group (CCEMG) and the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre). CCEMG - EPPI-Centre Cost Converter 2016 [Available from: https://eppi.ioe.ac.uk/costconversion/default.aspx.

88. Department of Health. Hospital and Community Health Services (HCHS) pay and price inflation 2016 [Available from: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=2ahUKEwiVudX3lNvfAhUJShUIHSBBBUAQFjABegQICBAC&url=http%3A%2F%2Fwww.info.doh.gov.uk%2Fdoh%2Ffinman.nsf%2Faf3d43e36a4c8f8500256722005b77f8%2F360a47827991d10a80258036002d8d9f%2F%24FILE%2F2015.16%2520Pay%2520%26%2520Price%2520series.xlsx&usg=AOvVaw0Ny2teATbwmkIyuljThctL.

89. NHS. Ensuring the Value of Clinical Commissioning Group (CCG) Commissioned Biologics Medicines – Biologics Value Tariff Agreement 2017 [Available from: http://www.lambethccg.nhs.uk/news-and-publications/meeting-papers/lambeth-borough-prescribing-committee/Lambeth%20Borough%20Prescribing%20Committee/Prescribing%20Rebates/Biologics%20Value%20Agreement%202017-19%20FINAL.pdf.

90. NHS England. Regional Medicines Optimisation Committee Briefing Best Value Biologicals: Adalimumab Update 3 2018 [Available from: https://www.sps.nhs.uk/wp-content/uploads/2018/07/Adalimumab-RMOC-Briefing-Final-July.pdf.

91. NHS England. Commissioning framework for biological medicines (including biosimilar medicines) 2017 [Available from: https://www.england.nhs.uk/wp-content/uploads/2017/09/biosimilar-medicines-commissioning-framework.pdf.

92. NHS England. Medicines Optimisation Dashboard [Available from: https://apps.nhsbsa.nhs.uk/MOD/AtlasTrustsMedsOp/atlas.html.

93. Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Annals of the Rheumatic Diseases. 2017;76(6):960-77.

94. Smolen JS, Nash P, Durez P, Hall S, Ilivanova E, Irazoque-Palazuelos F, et al. Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. Lancet. 2013;381(9870):918-29.

95. National Insitute of Health and Care Excellence (NICE). Tocilizumab for the treatment of rheumatoid arthritis (TA247). London: NICE; 2012.

96. Datapharm. The electronic Medicines Compendium: Remicade 100mg powder for concentrate for solution for infusion 2018 [Available from: https://www.medicines.org.uk/emc/product/3831/smpc.

97. National Institute for Health and Care Excellence (NICE). Therapeutic monitoring of TNF-alpha inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, IDKmonitor ELISA kits, and Promonitor ELISA kits) (DG22) 2016 [Available from: https://www.nice.org.uk/guidance/dg22.

98. National Institute for Health and Care Excellence (NICE). Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (TA329) 2015 [Available from: https://www.nice.org.uk/guidance/ta329.

99. Kobelt G, Lindgren P, Lindroth Y, Jacobson L, Eberhardt K. Modelling the effect of function and disease activity on costs and quality of life in rheumatoid arthritis. Rheumatology (Oxford). 2005;44(9):1169-75.

100. Burn E, Edwards CJ, Murray DW, Silman A, Cooper C, Arden NK, et al. Trends and determinants of length of stay and hospital reimbursement following knee and hip replacement: evidence from linked primary care and NHS hospital records from 1997 to 2014. BMJ Open. 2018;8(1):e019146.

101. Barnabe C, Thanh NX, Ohinmaa A, Homik J, Barr SG, Martin L, et al. Healthcare service utilisation costs are reduced when rheumatoid arthritis patients achieve sustained remission. Ann Rheum Dis. 2013;72(10):1664-8.

102. Beresniak A, Gossec L, Goupille P, Saraux A, Bamberger M, Bregman B, et al. Direct cost-modeling of rheumatoid arthritis according to disease activity categories in France. J Rheumatol. 2011;38(3):439-45.

103. Aletaha D, Funovits J, Keystone EC, Smolen JS. Disease activity early in the course of treatment predicts response to therapy after one year in rheumatoid arthritis patients. Arthritis Rheum. 2007;56(10):3226-35.

104. Pennington B, Davis S. Mapping from the Health Assessment Questionnaire to the EQ-5D: The Impact of Different Algorithms on Cost-Effectiveness Results. Value Health. 2014;17(8):762-71.

105. 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(6):963-71.

106. Kobelt G, Jonsson L, Lindgren P, Young A, Eberhardt K. Modeling the progression of rheumatoid arthritis: a two-country model to estimate costs and consequences of rheumatoid arthritis. Arthritis Rheum. 2002;46(9):2310-9.

107. 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(5):551-9.

108. Hernandez Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. Rheumatology (Oxford). 2013;52(5):944-50.

109. Alten R, Pohl C, Choy EH, Christensen R, Furst DE, Hewlett SE. Developing a construct to evaluate flares in rheumatoid arthritis: a conceptual report of the OMERACT RA flare definition working group. J Rheumatol. 2011;38.

110. Goh L, Jewell T, Laversuch C, Samanta A. A systematic review of the influence of anti-TNF on infection rates in patients with rheumatoid arthritis. Rev Bras Reumatol. 2013;53(6):501-15.

111. Listing J, Gerhold K, Zink A. The risk of infections associated with rheumatoid arthritis, with its comorbidity and treatment. Rheumatology (Oxford). 2013;52(1):53-61.

112. Gülfe A, Wallman JK, Kristensen LE. EuroQol-5 dimensions utility gain according to British and Swedish preference sets in rheumatoid arthritis treated with abatacept, rituximab, tocilizumab, or tumour necrosis factor inhibitors: a prospective cohort study from southern Sweden. Arthritis Research & Therapy. 2016;18(1):51.

113. Gulfe A, Kristensen LE, Saxne T, Jacobsson LT, Petersson IF, Geborek P. Rapid and sustained health utility gain in anti-tumour necrosis factor-treated inflammatory arthritis: observational data during 7 years in southern Sweden. Ann Rheum Dis. 2010;69(2):352-7.

114. Jani M, Chinoy H, Warren RB, Griffiths CE, Plant D, Fu B, et al. Clinical utility of random anti-tumor necrosis factor drug-level testing and measurement of antidrug antibodies on the long-term treatment response in rheumatoid arthritis. Arthritis Rheumatol. 2015;67(8):2011-9.

115. Jani M, Isaacs JD, Morgan AW, Wilson AG, Plant D, Hyrich K, et al. High frequency of antidrug antibodies and association of random drug levels with efficacy in certolizumab pegol-treated patients with rheumatoid arthritis: results from the BRAGGSS cohort. Ann Rheum Dis. 2017;76(1):208-13.

Appendix 1. Literature search strategies

ELISA for anti-TNF inhibitors in rheumatoid arthritis – clinical-effectiveness searches

Database: MEDLINE

Host: Ovid

Data Parameters: 1946 to July Week 2 2018

Date Searched: 20/7/2018

Searcher: SR

Hits: 1703

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.
2. anti* tumo?r* necrosis* factor*.tw.
3. Tumor Necrosis Factor-alpha/
4. (biologic* adj2 DMARD*).tw.
5. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.
6. ((disease modify* or disease-modify*) adj4 biologic*).tw.
7. exp Antibodies, Monoclonal/
8. anti* drug* antibod*.tw.
9. ADAb.tw.
10. etanercept.tw. or ETANERCEPT/
11. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.
12. (ETA or ETN).tw.
13. (enbrel or erelzi or benepali or lifmior or brenzys).tw.
14. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.
15. adalimumab.tw. or ADALIMUMAB/
16. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.
17. (ADA or ADL or ADM).tw.
18. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.
19. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.
20. infliximab.tw. or INFLIXIMAB/

21. (170277-31-3 or ta650 or ta 650 or ta-650).tw.

22. (INF or IFX).tw.

23. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.

24. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.

25. Certolizumab Pegol/ or certolizumab.tw.

26. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.

27. (CER or CZP).tw.

28. cimzia.tw.

29. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.

30. golimumab.tw.

31. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.

32. (GOL or GLM).tw.

33. simponi.tw.

34. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.

35. (biologic* adj2 agent*).tw.

36. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or
SB 5).tw.

37. (biosimilar* or (bio* adj1 similar*)).tw.

38. or/1-37

39. exp Enzyme-Linked Immunosorbent Assay/

40. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.

41. biohit healthcare.tw.

42. (proteomika* or *).tw.

43. (enzyme* adj3 immunoassay*).tw.

44. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.

45. ELISA*.tw.

46. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.

47. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.

48. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.

49. (mabtrack* or (mab adj3 track*) or mab-track*).tw.

50. sanquin.tw.

51. theradiag.tw.
52. (grifols or progenika).tw.
53. (r-biopharm or rbiopharm or r biopharm).tw.
54. ((drug* or trough) adj3 (level* or concentration)).tw.
55. or/39-54
56. exp Arthritis, Rheumatoid/
57. RA.tw.
58. Rheumarthrit*.tw.
59. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.
60. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.
61. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.
62. (Beauvais* adj2 disease*).tw.
63. or/56-62
64. 38 and 55 and 63
65. animals/ not humans/
66. 64 not 65

Database: MEDLINE In-Process & Other Non-Indexed Citations

Host: Ovid

Data Parameters: July 19 2018

Date Searched: 20/7/2018

Searcher: SR

Hits: 70

Database: EMBASE

Host: Ovid

Data Parameters: 1974 to 2018 July 19

Date Searched: 20/7/2018

Searcher: SR

Hits: 3807

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.
2. anti* tumo?r* necrosis* factor*.tw.
3. Tumor Necrosis Factor-alpha/
4. (biologic* adj2 DMARD*).tw.
5. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.
6. ((disease modify* or disease-modify*) adj4 biologic*).tw.
7. disease modifying antirheumatic drug/
8. monoclonal antibody/
9. anti* drug* antibod*.tw.
10. ADAb.tw.
11. etanercept.tw. or ETANERCEPT/
12. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.
13. (ETA or ETN).tw.
14. (enbrel or erelzi or benepali or lifmior or brenzys).tw.
15. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.
16. adalimumab.tw. or ADALIMUMAB/
17. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.
18. (ADA or ADL or ADM).tw.

19. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.

20. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.

21. infliximab.tw. or INFLIXIMAB/

22. (170277-31-3 or ta650 or ta 650 or ta-650).tw.

23. (INF or IFX).tw.

24. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.

25. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.

26. Certolizumab Pegol/ or certolizumab.tw.

27. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.

28. (CER or CZP).tw.

29. cimzia.tw.

30. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.

31. golimumab/ or golimumab.tw.

32. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.

33. (GOL or GLM).tw.

34. simponi.tw.

35. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.

36. (biologic* adj2 agent*).tw.

37. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5).tw.

38. biological product/ or biosimilar agent/

39. (biosimilar* or (bio* adj1 similar*)).tw.

40. or/1-39

41. exp Enzyme-Linked Immunosorbent Assay/

42. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.

43. biohit healthcare.tw.

44. (proteomika* or promonitor*).tw.

45. (enzyme* adj3 immunoassay*).tw.

46. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.

47. ELISA*.tw.

48. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.

49. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.

50. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.

51. (mabtrack* or (mab adj3 track*) or mab-track*).tw.

52. sanquin.tw.

53. theradiag.tw.

54. (grifols or progenika).tw.

55. (r-biopharm or rbiopharm or r biopharm).tw.

56. ((drug* or trough) adj3 (level* or concentration)).tw.

57. or/41-56

58. exp Arthritis, Rheumatoid/

59. RA.tw.

60. Rheumarthrit*.tw.

61. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.

62. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.

63. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.

64. (Beauvais* adj2 disease*).tw.

65. or/58-64

66. 40 and 57 and 65

67. (exp animal/ or nonhuman/) not exp human/

68. 66 not 67

Database: Web of Science (SCI and CPCI-S)

Host: Thomson Reuters

Data Parameters: n/a

Date Searched: 20/7/2018

Searcher: SR

Hits: 3633

Strategy:

#1 TS=(anti-TNF* or antiTNF* or (TNF* near/1 (inhibit* or block*))) OR TS=tumo$r* necrosis* factor* alpha OR TS= (biologic* near/1 DMARD*) OR TS=(biologic* near/3 antirheumati*) OR TS=(anti rheumati* near/3 biologic*) OR TS=(disease modify* near/3 biologic*) OR TS=anti* drug* antibod* OR TS=ADAb OR TS=anti* tumo$r* necrosis* factor* OR TS=monoclonal antibod*

#2 TS=etanercept OR TS=(tnr001 or tnr 001 or tnr-001 or 185243-69-0) OR TS=(ETA or ETN) OR TS=(enbrel or erelzi or benepali or lifmior or brenzys) OR TS=(anti-etanercept* or antietanercept* or anti near/2 etanercept*)

#3 TS=adalimumab OR TS=(d 2e7 or d2e7 or d-2e7 or 331731-18-1) OR TS=(ADA or ADL or ADM) OR TS=(humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz) OR TS=(anti-adalimumab* or antiadalimumab* or anti near/2 adalimumab*)

#4 TS= infliximab OR TS=(170277-31-3 or ta650 or ta 650 or ta-650) OR TS=(INF or IFX) OR TS=(anti-infliximab* or antiinfliximab* or anti near/2 infliximab*) OR TS=(remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi)

#5 TS=certolizumab OR TS=(cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2) OR TS=(CER or CZP) OR TS=cimzia OR TS=(anti-certolizumab* or anticertolizumab* or anti near/2 certolizumab*)

#6 TS=golimumab OR TS=(cnto 148 or cnto148 or cnto-148 or 476181-74-5) OR TS=(GOL or GLM) OR TS=simponi OR TS=(anti-golimumab* or antigolimumab* or anti near/2 golimumab*)

#7 TS=(biologic* near/1 agent*) OR TS=(CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5) OR TS=(biosimilar* or bio* similar*)

#8 #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1

#9 TS=(immundiagnostik* or immunodiagnostik* or immunediagnostik*) OR TS=biohit healthcare OR TS=(proteomika* or promonitor*) OR TS=(enzyme* near/2 immunoassay*) OR TS=(enzyme* near/2 immuno* assay*) OR TS=(enzyme* near/2 immuno* test*) OR TS=ELISA*

#10 TS= (idkmonitor* or idk near/2 monitor* or idk-monitor*) OR TS=(lisa near/2 tracker* or lisa-tracker* or lisatracker*) OR TS=(ridascreen* or rida near/2 screen* or rida-screen*) OR TS=(mabtrack* or mab near/2 track* or mab-track*) OR TS=(sanquin or theradiag) OR TS=(grifols or progenika) OR TS=(r-biopharm or rbiopharm or r biopharm) OR TS= ((drug* or trough) near/2 (level* or concentration))

#11 #10 OR #9

#12 TS=RA OR TS=Rheumarthrit* OR TS=((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) near/3 (arthrit* or arthros* or polyarthrit* or factor*)) OR TS=(chronic* near/3 polyarthrit*) OR TS=(chronic* near/3 poly arthrit*) OR TS=(chronic* near/3 rheumati*) OR TS=((Inflammat* or pain* or swell* or stiff*) near/3 (joint* or synovial*)) OR TS=(Beauvais* adj2 disease*)

#13 #12 AND #11 AND #8
Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2018;

Database: Cochrane Library

Host: Cochrane Collaboration

Data Parameters: CDSR Issue 7 of 12, July 2018; CENTRAL Issue 6 of 12, June 2018

Date Searched: 20/7/2018

Searcher: SR

Hits: 255

Strategy:

#1 (anti-TNF* or antiTNF* or (TNF* near/2 (inhibit* or block*))):ti,ab,kw

#2 "anti* tumo*r* necrosis* factor*":ti,ab,kw

#3 MeSH descriptor: [Tumor Necrosis Factor-alpha] this term only

#4 (biologic* near/2 DMARD*):ti,ab,kw

#5 ((antirheumati* or "anti rheumati*" or anti-rheumati*) near/4 biologic*):ti,ab,kw

#6 (("disease modify*" or disease-modify*) near/4 biologic*):ti,ab,kw

#7 MeSH descriptor: [Antibodies, Monoclonal] explode all trees

#8 "anti* drug* antibod*":ti,ab,kw

#9 ADAb:ti,ab

#10 etanercept:ti,ab,kw

#11 MeSH descriptor: [Etanercept] this term only

#12 (tnr001 or "tnr 001" or tnr-001 or 185243-69-0):ti,ab

#13 (ETA or ETN):ti,ab

#14 (enbrel or erelzi or benepali or lifmior or brenzys):ti,ab,kw

#15 (anti-etanercept* or antietanercept* or (anti near/3 etanercept*)):ti,ab,kw

#16 adalimumab:ti,ab,kw

#17 MeSH descriptor: [Adalimumab] this term only

#18 ("d 2e7" or d2e7 or d-2e7 or 331731-18-1):ti,ab

#19 (ADA or ADL or ADM):ti,ab

#20 (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz):ti,ab,kw

#21 (anti-adalimumab* or antiadalimumab* or (anti near/3 adalimumab*)):ti,ab,kw

#22 infliximab:ti,ab,kw

#23 MeSH descriptor: [Infliximab] this term only

#24 (170277-31-3 or ta650 or "ta 650" or ta-650):ti,ab

#25 (INF or IFX):ti,ab

#26 (anti-infliximab* or antiinfliximab* or (anti near/3 infliximab*)):ti,ab,kw

#27 (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi):ti,ab,kw

#28 certolizumab:ti,ab,kw

#29 MeSH descriptor: [Certolizumab Pegol] this term only

#30 (cdp870 or "cdp 870" or cdp-870 or 428863-50-7 or 1132819-27-2):ti,ab

#31 (CER or CZP):ti,ab

#32 cimzia:ti,ab,kw

#33 (anti-certolizumab* or anticertolizumab* or (anti near/3 certolizumab*)):ti,ab,kw

#34 golimumab:ti,ab,kw

#35 ("cnto 148" or cnto148 or cnto-148 or 476181-74-5):ti,ab

#36 (GOL or GLM):ti,ab

#37 simponi:ti,ab,kw

#38 (anti-golimumab* or antigolimumab* or (anti near/3 golimumab*)):ti,ab,kw

#39 (biologic* near/2 agent*):ti,ab,kw

#40 (CT-P13 or CTP13 or "CT P13" or SB2 or SB-2 or "SB 2" or SB4 or SB-4 or "SB 4" or SB-5 or SB5 or "SB 5"):ti,ab

#41 (biosimilar* or "bio* similar*"):ti,ab,kw

#42 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41

#43 MeSH descriptor: [Enzyme-Linked Immunosorbent Assay] explode all trees

#44 (immundiagnostik* or immunodiagnostik* or immunediagnostik*):ti,ab,kw

#45 "biohit healthcare":ti,ab,kw

#46 (proteomika* or promonitor*):ti,ab,kw

#47 (enzyme* near/3 immunoassay*):ti,ab,kw

#48 (enzyme* near/3 ("immuno* assay*" or "immuno* test*")):ti,ab,kw

#49 ELISA*:ti,ab,kw

#50 (idkmonitor* or (idk near/3 monitor*) or idk-monitor*):ti,ab,kw

#51 ((lisa near/3 tracker*) or lisa-tracker* or lisatracker*):ti,ab,kw

#52 (ridascreen* or (rida near/3 screen*) or rida-screen*):ti,ab,kw

#53 (mabtrack* or (mab near/3 track*) or mab-track*):ti,ab,kw

#54 (sanquin or theradiag):ti,ab,kw

#55 (grifols or progenika):ti,ab,kw

#56 (r-biopharm or rbiopharm or "r biopharm"):ti,ab,kw

#57 ((drug* or trough) near/3 (level* or concentration)):ti,ab,kw

#58 #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57

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

#60 RA:ti,ab

#61 Rheumarthrit*.ti,ab,kw

#62 ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) near/4 (arthrit* or arthros* or polyarthrit* or factor*)):ti,ab,kw

#63 (Chronic* near/4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)):ti,ab,kw

#64 ((Inflammat* or pain* or swell* or stiff*) near/4 (joint* or synovial*)):ti,ab,kw

#65 (Beauvais* near/2 disease*):ti,ab,kw

#66 #59 or #60 or #61 or #62 or #64 or #65

#67 #42 and #58 and #66

Number of hits per database and in total

Database	Hits
MEDLINE	1,703
MEDLINE In-Process	70
EMBASE	3,807
Web of Science (SCI and SCCI)	3,633
Cochrane	255
Total records	**9,468**
Duplicates	**2,851**
Total unique records	**6,617**

Backward citation chasing

Citation chasing yielded 42 further references (after de-duplicating and checking against already screened papers), on 12 September 2018.

ELISA for anti-TNF inhibitors in rheumatoid arthritis – cost-effectiveness searches

Database: MEDLINE

Host: Ovid

Data Parameters: 1946 to July Week 2 2018

Date Searched: 26/7/2018

Searcher: SR

Hits: 4

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.
2. anti* tumo?r* necrosis* factor*.tw.
3. Tumor Necrosis Factor-alpha/
4. (biologic* adj2 DMARD*).tw.
5. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.
6. ((disease modify* or disease-modify*) adj4 biologic*).tw.
7. exp Antibodies, Monoclonal/
8. anti* drug* antibod*.tw.
9. ADAb.tw.
10. etanercept.tw. or ETANERCEPT/
11. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.
12. (ETA or ETN).tw.
13. (enbrel or erelzi or benepali or lifmior or brenzys).tw.
14. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.
15. adalimumab.tw. or ADALIMUMAB/
16. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.
17. (ADA or ADL or ADM).tw.
18. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.
19. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.
20. infliximab.tw. or INFLIXIMAB/
21. (170277-31-3 or ta650 or ta 650 or ta-650).tw.
22. (INF or IFX).tw.

23. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.

24. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.

25. Certolizumab Pegol/ or certolizumab.tw.

26. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.

27. (CER or CZP).tw.

28. cimzia.tw.

29. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.

30. golimumab.tw.

31. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.

32. (GOL or GLM).tw.

33. simponi.tw.

34. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.

35. (biologic* adj2 agent*).tw.

36. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5).tw.

37. (biosimilar* or (bio* adj1 similar*)).tw.

38. or/1-37

39. exp Enzyme-Linked Immunosorbent Assay/

40. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.

41. biohit healthcare.tw.

42. (proteomika* or promonitor*).tw.

43. (enzyme* adj3 immunoassay*).tw.

44. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.

45. ELISA*.tw.

46. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.

47. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.

48. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.

49. (mabtrack* or (mab adj3 track*) or mab-track*).tw.

50. sanquin.tw.

51. theradiag.tw.

52. (grifols or progenika).tw.

53. (r-biopharm or rbiopharm or r biopharm).tw.

54. ((drug* or trough) adj3 (level* or concentration)).tw.

55. or/39-54

56. exp Arthritis, Rheumatoid/

57. RA.tw.

58. Rheumarthrit*.tw.

59. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.

60. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.

61. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.

62. (Beauvais* adj2 disease*).tw.

63. or/56-62

64. 38 and 55 and 63

65. animals/ not humans/

66. 64 not 65

67. Economics/

68. exp "Costs and Cost Analysis"/

69. Economics, Nursing/

70. Economics, Medical/

71. Economics, Pharmaceutical/

72. exp Economics, Hospital/

73. Economics, Dental/

74. exp "Fees and Charges"/

75. exp Budgets/

76. budget*.ti,ab,kf.

77. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kf.

78. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2

79. (cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kf.

80. (value adj2 (money or monetary)).ti,ab,kf.

81. exp models, economic/

82. economic model*.ab,kf.
83. markov chains/
84. markov.ti,ab,kf.
85. monte carlo method/
86. monte carlo.ti,ab,kf.
87. exp Decision Theory/
88. (decision* adj2 (tree* or analy* or model*)).ti,ab,kf.
89. or/67-88
90. 66 and 89

Database: MEDLINE In-Process & Other Non-Indexed Citations

Host: Ovid

Data Parameters: July 25 2018

Date Searched: 25/7/2018

Searcher: SR

Hits: 1

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.
2. anti* tumo?r* necrosis* factor*.tw.
3. (biologic* adj2 DMARD*).tw.
4. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.
5. ((disease modify* or disease-modify*) adj4 biologic*).tw.
6. anti* drug* antibod*.tw.
7. ADAb.tw.
8. etanercept.tw.
9. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.
10. (ETA or ETN).tw.
11. (enbrel or erelzi or benepali or lifmior or brenzys).tw.
12. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.

13. adalimumab.tw.

14. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.

15. (ADA or ADL or ADM).tw.

16. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.

17. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.

18. infliximab.tw.

19. (170277-31-3 or ta650 or ta 650 or ta-650).tw.

20. (INF or IFX).tw.

21. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.

22. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.

23. certolizumab.tw.

24. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.

25. (CER or CZP).tw.

26. cimzia.tw.

27. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.

28. golimumab.tw.

29. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.

30. (GOL or GLM).tw.

31. simponi.tw.

32. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.

33. (biologic* adj2 agent*).tw.

34. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5).tw.

35. (biosimilar* or (bio* adj1 similar*)).tw.

36. or/1-35

37. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.

38. biohit healthcare.tw.

39. (proteomika* or promonitor*).tw.

40. (enzyme* adj3 immunoassay*).tw.

41. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.

42. ELISA*.tw.

43. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.

44. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.

45. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.

46. (mabtrack* or (mab adj3 track*) or mab-track*).tw.

47. sanquin.tw.

48. theradiag.tw.

49. (grifols or progenika).tw.

50. (r-biopharm or rbiopharm or r biopharm).tw.

51. ((drug* or trough) adj3 (level* or concentration)).tw.

52. or/37-51

53. RA.tw.

54. Rheumarthrit*.tw.

55. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.

56. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.

57. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.

58. (Beauvais* adj2 disease*).tw.

59. or/53-58

60. 36 and 52 and 59

61. budget*.ti,ab,kf.

62. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kf.

63. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2

64. (cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kf.

65. (value adj2 (money or monetary)).ti,ab,kf.

66. economic model*.ab,kf.

67. markov.ti,ab,kf.

68. monte carlo.ti,ab,kf.

69. (decision* adj2 (tree* or analy* or model*)).ti,ab,kf.

70. or/61-69

71. 60 and 70

Database: EMBASE

Host: Ovid

Data Parameters: 1974 to 2018 July 25

Date Searched: 26/7/2018

Searcher: SR

Hits: 102

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.

2. anti* tumo?r* necrosis* factor*.tw.

3. Tumor Necrosis Factor-alpha/

4. (biologic* adj2 DMARD*).tw.

5. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.

6. ((disease modify* or disease-modify*) adj4 biologic*).tw.

7. disease modifying antirheumatic drug/

8. monoclonal antibody/

9. anti* drug* antibod*.tw.

10. ADAb.tw.

11. etanercept.tw. or ETANERCEPT/

12. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.

13. (ETA or ETN).tw.

14. (enbrel or erelzi or benepali or lifmior or brenzys).tw.

15. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.

16. adalimumab.tw. or ADALIMUMAB/

17. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.
18. (ADA or ADL or ADM).tw.
19. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.
20. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.
21. infliximab.tw. or INFLIXIMAB/
22. (170277-31-3 or ta650 or ta 650 or ta-650).tw.
23. (INF or IFX).tw.
24. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.
25. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.
26. Certolizumab Pegol/ or certolizumab.tw.
27. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.
28. (CER or CZP).tw.
29. cimzia.tw.
30. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.
31. golimumab/ or golimumab.tw.
32. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.
33. (GOL or GLM).tw.
34. simponi.tw.
35. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.
36. (biologic* adj2 agent*).tw.
37. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5).tw.
38. biological product/ or biosimilar agent/
39. (biosimilar* or (bio* adj1 similar*)).tw.
40. or/1-39
41. exp Enzyme-Linked Immunosorbent Assay/
42. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.
43. biohit healthcare.tw.
44. (proteomika* or promonitor*).tw.

45. (enzyme* adj3 immunoassay*).tw.
46. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.
47. ELISA*.tw.
48. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.
49. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.
50. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.
51. (mabtrack* or (mab adj3 track*) or mab-track*).tw.
52. sanquin.tw.
53. theradiag.tw.
54. (grifols or progenika).tw.
55. (r-biopharm or rbiopharm or r biopharm).tw.
56. ((drug* or trough) adj3 (level* or concentration)).tw.
57. or/41-56
58. exp Arthritis, Rheumatoid/
59. RA.tw.
60. Rheumarthrit*.tw.
61. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.
62. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.
63. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.
64. (Beauvais* adj2 disease*).tw.
65. or/58-64
66. 40 and 57 and 65
67. (exp animal/ or nonhuman/) not exp human/
68. 66 not 67
69. Economics/
70. Cost/
71. exp Health Economics/
72. Budget/

73. budget*.ti,ab,kw.

74. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kw.

75. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ab. /freq=2

76. (cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kw.

77. (value adj2 (money or monetary)).ti,ab,kw.

78. Statistical Model/

79. economic model*.ab,kw.

80. Probability/

81. markov.ti,ab,kw.

82. monte carlo method/

83. monte carlo.ti,ab,kw.

84. Decision Theory/

85. Decision Tree/

86. (decision* adj2 (tree* or analy* or model*)).ti,ab,kw.

87. or/69-86

88. 68 and 87

Database: Web of Science (SCI and CPCI-S)

Host: Thomson Reuters

Data Parameters: n/a

Date Searched: 2/7/2018

Searcher: SR

Hits: 63

Strategy:

1. #1 TS=(anti-TNF* or antiTNF* or (TNF* near/1 (inhibit* or block*))) OR TS=tumo$r* necrosis* factor* alpha OR TS= (biologic* near/1 DMARD*) OR TS=(biologic* near/3

antirheumati*) OR TS=(anti rheumati* near/3 biologic*) OR TS=(disease modify* near/3 biologic*) OR TS=anti* drug* antibod* OR TS=ADAb OR TS=anti* tumo$r* necrosis* factor* OR TS=monoclonal antibod*

2. #2 TS=etanercept OR TS=(tnr001 or tnr 001 or tnr-001 or 185243-69-0) OR TS=(ETA or ETN) OR TS=(enbrel or erelzi or benepali or lifmior or brenzys) OR TS=(anti-etanercept* or antietanercept* or anti near/2 etanercept*)
3. #3 TS=adalimumab OR TS=(d 2e7 or d2e7 or d-2e7 or 331731-18-1) OR TS=(ADA or ADL or ADM) OR TS=(humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz) OR TS=(anti-adalimumab* or antiadalimumab* or anti near/2 adalimumab*)
4. #4 TS= infliximab OR TS=(170277-31-3 or ta650 or ta 650 or ta-650) OR TS=(INF or IFX) OR TS=(anti-infliximab* or antiinfliximab* or anti near/2 infliximab*) OR TS=(remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi)
5. #5 TS=certolizumab OR TS=(cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2) OR TS=(CER or CZP) OR TS=cimzia OR TS=(anti-certolizumab* or anticertolizumab* or anti near/2 certolizumab*)
6. #6 TS=golimumab OR TS=(cnto 148 or cnto148 or cnto-148 or 476181-74-5) OR TS=(GOL or GLM) OR TS=simponi OR TS=(anti-golimumab* or antigolimumab* or anti near/2 golimumab*)
7. #7 TS=(biologic* near/1 agent*) OR TS=(CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5) OR TS=(biosimilar* or bio* similar*)
8. #8 #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1
9. #9 TS=(immundiagnostik* or immunodiagnostik* or immunediagnostik*) OR TS=biohit healthcare OR TS=(proteomika* or promonitor*) OR TS=(enzyme* near/2 immunoassay*) OR TS=(enzyme* near/2 immuno* assay*) OR TS=(enzyme* near/2 immuno* test*) OR TS=ELISA*
10. #10 TS= (idkmonitor* or idk near/2 monitor* or idk-monitor*) OR TS=(lisa near/2 tracker* or lisa-tracker* or lisatracker*) OR TS=(ridascreen* or rida near/2 screen* or rida-screen*) OR TS=(mabtrack* or mab near/2 track* or mab-track*) OR TS=(sanquin or theradiag) OR TS=(grifols or progenika) OR TS=(r-biopharm or rbiopharm or r biopharm) OR TS= ((drug* or trough) near/2 (level* or concentration))
11. #11 #10 OR #9

12. #12 TS=RA OR TS=Rheumarthrit* OR TS=((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) near/3 (arthrit* or arthros* or polyarthrit* or factor*)) OR TS=(chronic* near/3 polyarthrit*) OR TS=(chronic* near/3 poly arthrit*) OR TS=(chronic* near/3 rheumati*) OR TS=((Inflammat* or pain* or swell* or stiff*) near/3 (joint* or synovial*)) OR TS=(Beauvais* adj2 disease*)

13. #13 #12 AND #11 AND #8

14. TS=((pharmacoeconomic* or socioeconomics or economic* or pric* or cost* or cba or cea or cua or "health utilit*" or "value for money"))

15. #14 and #15

Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2018;

Database: NHS EED

Host: Cochrane Library

Data Parameters: Issue 2 of 4 April 2015

Date Searched: 26/7/2018

Searcher: SR

Hits: 0

Strategy:

#1 (anti-TNF* or antiTNF* or (TNF* near/2 (inhibit* or block*))):ti,ab,kw

#2 "anti* tumo*r* necrosis* factor*":ti,ab,kw

#3 MeSH descriptor: [Tumor Necrosis Factor-alpha] this term only

#4 (biologic* near/2 DMARD*):ti,ab,kw

#5 ((antirheumati* or "anti rheumati*" or anti-rheumati*) near/4 biologic*):ti,ab,kw

#6 (("disease modify*" or disease-modify*) near/4 biologic*):ti,ab,kw

#7 MeSH descriptor: [Antibodies, Monoclonal] explode all trees

#8 "anti* drug* antibod*":ti,ab,kw

#9 ADAb:ti,ab

#10 etanercept:ti,ab,kw

#11 MeSH descriptor: [Etanercept] this term only

#12 (tnr001 or "tnr 001" or tnr-001 or 185243-69-0):ti,ab

#13 (ETA or ETN):ti,ab

#14 (enbrel or erelzi or benepali or lifmior or brenzys):ti,ab,kw

#15 (anti-etanercept* or antietanercept* or (anti near/3 etanercept*)):ti,ab,kw

#16 adalimumab:ti,ab,kw

#17 MeSH descriptor: [Adalimumab] this term only

#18 ("d 2e7" or d2e7 or d-2e7 or 331731-18-1):ti,ab

#19 (ADA or ADL or ADM):ti,ab

#20 (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz):ti,ab,kw

#21 (anti-adalimumab* or antiadalimumab* or (anti near/3 adalimumab*)):ti,ab,kw

#22 infliximab:ti,ab,kw

#23 MeSH descriptor: [Infliximab] this term only

#24 (170277-31-3 or ta650 or "ta 650" or ta-650):ti,ab

#25 (INF or IFX):ti,ab

#26 (anti-infliximab* or antiinfliximab* or (anti near/3 infliximab*)):ti,ab,kw

#27 (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi):ti,ab,kw

#28 certolizumab:ti,ab,kw

#29 MeSH descriptor: [Certolizumab Pegol] this term only

#30 (cdp870 or "cdp 870" or cdp-870 or 428863-50-7 or 1132819-27-2):ti,ab

#31 (CER or CZP):ti,ab

#32 cimzia:ti,ab,kw

#33 (anti-certolizumab* or anticertolizumab* or (anti near/3 certolizumab*)):ti,ab,kw

#34 golimumab:ti,ab,kw

#35 ("cnto 148" or cnto148 or cnto-148 or 476181-74-5):ti,ab

#36 (GOL or GLM):ti,ab

#37 simponi:ti,ab,kw

#38 (anti-golimumab* or antigolimumab* or (anti near/3 golimumab*)):ti,ab,kw

#39 (biologic* near/2 agent*):ti,ab,kw

#40 (CT-P13 or CTP13 or "CT P13" or SB2 or SB-2 or "SB 2" or SB4 or SB-4 or "SB 4" or SB-5 or SB5 or "SB 5"):ti,ab

#41 (biosimilar* or "bio* similar*"):ti,ab,kw

#42 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41

#43 MeSH descriptor: [Enzyme-Linked Immunosorbent Assay] explode all trees

#44 (immundiagnostik* or immunodiagnostik* or immunediagnostik*):ti,ab,kw

#45 "biohit healthcare":ti,ab,kw

#46 (proteomika* or promonitor*):ti,ab,kw

#47 (enzyme* near/3 immunoassay*):ti,ab,kw

#48 (enzyme* near/3 ("immuno* assay*" or "immuno* test*")):ti,ab,kw

#49 ELISA*:ti,ab,kw

#50 (idkmonitor* or (idk near/3 monitor*) or idk-monitor*):ti,ab,kw

#51 ((lisa near/3 tracker*) or lisa-tracker* or lisatracker*):ti,ab,kw

#52 (ridascreen* or (rida near/3 screen*) or rida-screen*):ti,ab,kw

#53 (mabtrack* or (mab near/3 track*) or mab-track*):ti,ab,kw

#54 (sanquin or theradiag):ti,ab,kw

#55 (grifols or progenika):ti,ab,kw

#56 (r-biopharm or rbiopharm or "r biopharm"):ti,ab,kw

#57 ((drug* or trough) near/3 (level* or concentration)):ti,ab,kw

#58 #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57

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

#60 RA:ti,ab

#61 Rheumarthrit*.ti,ab,kw

#62 ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) near/4 (arthrit* or arthros* or polyarthrit* or factor*)):ti,ab,kw

#63 (Chronic* near/4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)):ti,ab,kw

#64 ((Inflammat* or pain* or swell* or stiff*) near/4 (joint* or synovial*)):ti,ab,kw

#65 (Beauvais* near/2 disease*):ti,ab,kw

#66 #59 or #60 or #61 or #62 or #64 or #65

#67 #42 and #58 and #66

Database: EconLit

Host: EBSCO

Data Parameters: n/a

Date Searched: 2/7/2018

Searcher: SR

Hits: 56

Strategy:

1. TX Rheumarthrit*
2. TX ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) N4 (arthrit* or arthros* or polyarthrit* or factor*))
3. TX ((Chronic* N4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*))
4. TX ((Inflammat* or pain* or swell* or stiff*) N4 (joint* or synovial*))
5. S1 OR S2 OR S3 OR S4

Number of hits per database and in total

Database	Hits
MEDLINE	5
MEDLINE In-Process	1
EMBASE	102
Web of Science (SCI and SCCI)	63
Cochrane – HTA and NHS EED	0

Database	Hits
Econlit	56
Total records	**227**
Duplicates	**13**
Total unique records	**214**

ELISA for anti-TNF inhibitors in rheumatoid arthritis – utilities searches

Database: MEDLINE

Host: Ovid

Data Parameters: 1946 to July Week 3 2018

Date Searched: 30/7/2018

Searcher: SR

Hits: 136

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.
2. anti* tumo?r* necrosis* factor*.tw.
3. Tumor Necrosis Factor-alpha/
4. (biologic* adj2 DMARD*).tw.
5. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.
6. ((disease modify* or disease-modify*) adj4 biologic*).tw.
7. exp Antibodies, Monoclonal/
8. anti* drug* antibod*.tw.
9. ADAb.tw.
10. etanercept.tw. or ETANERCEPT/
11. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.
12. (ETA or ETN).tw.
13. (enbrel or erelzi or benepali or lifmior or brenzys).tw.
14. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.
15. adalimumab.tw. or ADALIMUMAB/

16. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.
17. (ADA or ADL or ADM).tw.
18. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.
19. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.
20. infliximab.tw. or INFLIXIMAB/
21. (170277-31-3 or ta650 or ta 650 or ta-650).tw.
22. (INF or IFX).tw.
23. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.
24. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.
25. Certolizumab Pegol/ or certolizumab.tw.
26. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.
27. (CER or CZP).tw.
28. cimzia.tw.
29. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.
30. golimumab.tw.
31. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.
32. (GOL or GLM).tw.
33. simponi.tw.
34. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.
35. (biologic* adj2 agent*).tw.
36. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5).tw.
37. (biosimilar* or (bio* adj1 similar*)).tw.
38. or/1-37
39. exp Enzyme-Linked Immunosorbent Assay/
40. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.
41. biohit healthcare.tw.
42. (proteomika* or promonitor*).tw.
43. (enzyme* adj3 immunoassay*).tw.

44. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.

45. ELISA*.tw.

46. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.

47. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.

48. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.

49. (mabtrack* or (mab adj3 track*) or mab-track*).tw.

50. sanquin.tw.

51. theradiag.tw.

52. (grifols or progenika).tw.

53. (r-biopharm or rbiopharm or r biopharm).tw.

54. ((drug* or trough) adj3 (level* or concentration)).tw.

55. or/39-54

56. exp Arthritis, Rheumatoid/

57. RA.tw.

58. Rheumarthrit*.tw.

59. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.

60. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.

61. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.

62. (Beauvais* adj2 disease*).tw.

63. or/56-62

64. 38 and 55 and 63

65. animals/ not humans/

66. 64 not 65

67. "Value of Life"/

68. Quality of Life/

69. quality of life.ti,kf.

70. ((instrument or instruments) adj3 quality of life).ab.

71. Quality-Adjusted Life Years/

72. quality adjusted life.ti,ab,kf.

73. (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kf.
74. disability adjusted life.ti,ab,kf.
75. daly*.ti,ab,kf.
76. (sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab,kf.
77. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6).ti,ab,kf.
78. (sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or shortform eight or short form eight).ti,ab,kf.
79. (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,kf.
80. (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab,kf.
81. (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,kf.
82. (hql or hqol or h qol or hrqol or hr qol).ti,ab,kf.
83. (hye or hyes).ti,ab,kf.
84. (health* adj2 year* adj2 equivalent*).ti,ab,kf.
85. (pqol or qls).ti,ab,kf.
86. (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb).ti,ab,kf.
87. nottingham health profile*.ti,ab,kf.
88. sickness impact profile.ti,ab,kf.
89. exp health status indicators/
90. (health adj3 (utilit* or status)).ti,ab,kf.
91. (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)).ti,ab,kf.
92. (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)).ti,ab,kf.
93. disutilit*.ti,ab,kf.
94. rosser.ti,ab,kf.
95. willingness to pay.ti,ab,kf.
96. standard gamble*.ti,ab,kf.
97. (time trade off or time tradeoff).ti,ab,kf.
98. tto.ti,ab,kf.

99. (hui or hui1 or hui2 or hui3).ti,ab,kf.
100. (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual).ti,ab,kf.
101. duke health profile.ti,ab,kf.
102. functional status questionnaire.ti,ab,kf.
103. dartmouth coop functional health assessment*.ti,ab,kf.
104. or/67-103
105. 66 and 104

Database: MEDLINE In-Process & Other Non-Indexed Citations

Host: Ovid

Data Parameters: July 27 2018

Date Searched: 30/7/2018

Searcher: SR

Hits: 2

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.
2. anti* tumo?r* necrosis* factor*.tw.
3. (biologic* adj2 DMARD*).tw.
4. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.
5. ((disease modify* or disease-modify*) adj4 biologic*).tw.
6. anti* drug* antibod*.tw.
7. ADAb.tw.
8. etanercept.tw.
9. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.
10. (ETA or ETN).tw.
11. (enbrel or erelzi or benepali or lifmior or brenzys).tw.
12. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.
13. adalimumab.tw.
14. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.

15. (ADA or ADL or ADM).tw.

16. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.

17. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.

18. infliximab.tw.

19. (170277-31-3 or ta650 or ta 650 or ta-650).tw.

20. (INF or IFX).tw.

21. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.

22. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.

23. certolizumab.tw.

24. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.

25. (CER or CZP).tw.

26. cimzia.tw.

27. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.

28. golimumab.tw.

29. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.

30. (GOL or GLM).tw.

31. simponi.tw.

32. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.

33. (biologic* adj2 agent*).tw.

34. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5).tw.

35. (biosimilar* or (bio* adj1 similar*)).tw.

36. or/1-35

37. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.

38. biohit healthcare.tw.

39. (proteomika* or promonitor*).tw.

40. (enzyme* adj3 immunoassay*).tw.

41. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.

42. ELISA*.tw.

43. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.

44. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.

45. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.

46. (mabtrack* or (mab adj3 track*) or mab-track*).tw.

47. sanquin.tw.

48. theradiag.tw.

49. (grifols or progenika).tw.

50. (r-biopharm or rbiopharm or r biopharm).tw.

51. ((drug* or trough) adj3 (level* or concentration)).tw.

52. or/37-51

53. RA.tw.

54. Rheumarthrit*.tw.

55. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.

56. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.

57. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.

58. (Beauvais* adj2 disease*).tw.

59. or/53-58

60. 36 and 52 and 59

61. quality of life.ti,kf.

70. ((instrument or instruments) adj3 quality of life).ab.

71. quality adjusted life.ti,ab,kf.

72. (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kf.

73. disability adjusted life.ti,ab,kf.

74. daly*.ti,ab,kf.

75. (sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab,kf.

76. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6).ti,ab,kf.

77. (sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or shortform eight or short form eight).ti,ab,kf.

78. (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,kf.

79. (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab,kf.

80. (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,kf.

81. (hql or hqol or h qol or hrqol or hr qol).ti,ab,kf.

82. (hye or hyes).ti,ab,kf.

83. (health* adj2 year* adj2 equivalent*).ti,ab,kf.

84. (pqol or qls).ti,ab,kf.

85. (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb).ti,ab,kf.

86. nottingham health profile*.ti,ab,kf.

87. sickness impact profile.ti,ab,kf.

88. (health adj3 (utilit* or status)).ti,ab,kf.

89. (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)).ti,ab,kf.

90. (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)).ti,ab,kf.

91. disutilit*.ti,ab,kf.

92. rosser.ti,ab,kf.

93. willingness to pay.ti,ab,kf.

94. standard gamble*.ti,ab,kf.

95. (time trade off or time tradeoff).ti,ab,kf.

96. tto.ti,ab,kf.

97. (hui or hui1 or hui2 or hui3).ti,ab,kf.

98. (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual).ti,ab,kf.

99. duke health profile.ti,ab,kf.

100. functional status questionnaire.ti,ab,kf.

101. dartmouth coop functional health assessment*.ti,ab,kf.

102. or/61-101

103. 60 and 102

Database: EMBASE

Host: Ovid

Data Parameters: 1974 to 2018 July 27

Date Searched: 30/7/2018

Searcher: SR

Hits: 64

Strategy:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.
2. anti* tumo?r* necrosis* factor*.tw.
3. Tumor Necrosis Factor-alpha/
4. (biologic* adj2 DMARD*).tw.
5. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.
6. ((disease modify* or disease-modify*) adj4 biologic*).tw.
7. disease modifying antirheumatic drug/
8. monoclonal antibody/
9. anti* drug* antibod*.tw.
10. ADAb.tw.
11. etanercept.tw. or ETANERCEPT/
12. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.
13. (ETA or ETN).tw.
14. (enbrel or erelzi or benepali or lifmior or brenzys).tw.
15. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.
16. adalimumab.tw. or ADALIMUMAB/
17. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.
18. (ADA or ADL or ADM).tw.
19. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.

20. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.

21. infliximab.tw. or INFLIXIMAB/

22. (170277-31-3 or ta650 or ta 650 or ta-650).tw.

23. (INF or IFX).tw.

24. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.

25. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.

26. Certolizumab Pegol/ or certolizumab.tw.

27. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.

28. (CER or CZP).tw.

29. cimzia.tw.

30. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.

31. golimumab/ or golimumab.tw.

32. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.

33. (GOL or GLM).tw.

34. simponi.tw.

35. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.

36. (biologic* adj2 agent*).tw.

37. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or
SB 5).tw.

38. biological product/ or biosimilar agent/

39. (biosimilar* or (bio* adj1 similar*)).tw.

40. or/1-39

41. exp Enzyme-Linked Immunosorbent Assay/

42. (immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.

43. biohit healthcare.tw.

44. (proteomika* or promonitor*).tw.

45. (enzyme* adj3 immunoassay*).tw.

46. (enzyme* adj3 (immuno* assay* or immuno* test*)).tw.

47. ELISA*.tw.

48. (idkmonitor* or (idk adj3 monitor*) or idk-monitor*).tw.
49. ((lisa adj3 tracker*) or lisa-tracker* or lisatracker*).tw.
50. (ridascreen* or (rida adj3 screen*) or rida-screen*).tw.
51. (mabtrack* or (mab adj3 track*) or mab-track*).tw.
52. sanquin.tw.
53. theradiag.tw.
54. (grifols or progenika).tw.
55. (r-biopharm or rbiopharm or r biopharm).tw.
56. ((drug* or trough) adj3 (level* or concentration)).tw.
57. or/41-56
58. exp Arthritis, Rheumatoid/
59. RA.tw.
60. Rheumarthrit*.tw.
61. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.
62. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.
63. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.
64. (Beauvais* adj2 disease*).tw.
65. or/58-64
66. 40 and 57 and 65
67. (exp animal/ or nonhuman/) not exp human/
68. 66 not 67
69. socioeconomics/
70. exp Quality of Life/
71. quality of life.ti,kw.
72. ((instrument or instruments) adj3 quality of life).ab.
73. Quality-Adjusted Life Year/
74. quality adjusted life.ti,ab,kw.

75. (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kw.

76. disability adjusted life.ti,ab,kw.

77. daly*.ti,ab,kw.

78. (sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab,kw.

79. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6).ti,ab,kw.

80. (sf8 or sf 8 or sf eight or sfeight or shortform 8 or shortform 8 or shortform8 or short form8 or shortform eight or short form eight).ti,ab,kw.

81. (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,kw.

82. (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab,kw.

83. (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,kw.

84. (hql or hqol or h qol or hrqol or hr qol).ti,ab,kw.

85. (hye or hyes).ti,ab,kw.

86. (health* adj2 year* adj2 equivalent*).ti,ab,kw.

87. (pqol or qls).ti,ab,kw.

88. (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb).ti,ab,kw.

89. nottingham health profile*.ti,ab,kw.

90. nottingham health profile/

91. sickness impact profile.ti,ab,kw.

92. sickness impact profile/

93. health status indicator/

94. (health adj3 (utilit* or status)).ti,ab,kw.

95. (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)).ti,ab,kw.

96. (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)).ti,ab,kw.

97. disutilit*.ti,ab,kw.

98. rosser.ti,ab,kw.

99. willingness to pay.ti,ab,kw.

100. standard gamble*.ti,ab,kw.

101. (time trade off or time tradeoff).ti,ab,kw.

102. tto.ti,ab,kw.

103. (hui or hui1 or hui2 or hui3).ti,ab,kw.

104. (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual).ti,ab,kw.

105. duke health profile.ti,ab,kw.

106. functional status questionnaire.ti,ab,kw.

107. dartmouth coop functional health assessment*.ti,ab,kw.

108. or/67-107

109. 68 and 108

Database: NHS EED

Host: Cochrane Library

Data Parameters: Issue 2 of 4 April 2015

Date Searched: 30/7/2018

Searcher: SR

Hits: 0

Strategy:

#1 (anti-TNF* or antiTNF* or (TNF* near/2 (inhibit* or block*))):ti,ab,kw

#2 "anti* tumo*r* necrosis* factor*":ti,ab,kw

#3 MeSH descriptor: [Tumor Necrosis Factor-alpha] this term only

#4 (biologic* near/2 DMARD*):ti,ab,kw

#5 ((antirheumati* or "anti rheumati*" or anti-rheumati*) near/4 biologic*):ti,ab,kw

#6 (("disease modify*" or disease-modify*) near/4 biologic*):ti,ab,kw

#7 MeSH descriptor: [Antibodies, Monoclonal] explode all trees

#8 "anti* drug* antibod*":ti,ab,kw

#9 ADAb:ti,ab

#10 etanercept:ti,ab,kw

#11 MeSH descriptor: [Etanercept] this term only

#12 (tnr001 or "tnr 001" or tnr-001 or 185243-69-0):ti,ab

#13 (ETA or ETN):ti,ab

#14 (enbrel or erelzi or benepali or lifmior or brenzys):ti,ab,kw

#15 (anti-etanercept* or antietanercept* or (anti near/3 etanercept*)):ti,ab,kw

#16 adalimumab:ti,ab,kw

#17 MeSH descriptor: [Adalimumab] this term only

#18 ("d 2e7" or d2e7 or d-2e7 or 331731-18-1):ti,ab

#19 (ADA or ADL or ADM):ti,ab

#20 (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz):ti,ab,kw

#21 (anti-adalimumab* or antiadalimumab* or (anti near/3 adalimumab*)):ti,ab,kw

#22 infliximab:ti,ab,kw

#23 MeSH descriptor: [Infliximab] this term only

#24 (170277-31-3 or ta650 or "ta 650" or ta-650):ti,ab

#25 (INF or IFX):ti,ab

#26 (anti-infliximab* or antiinfliximab* or (anti near/3 infliximab*)):ti,ab,kw

#27 (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi):ti,ab,kw

#28 certolizumab:ti,ab,kw

#29 MeSH descriptor: [Certolizumab Pegol] this term only

#30 (cdp870 or "cdp 870" or cdp-870 or 428863-50-7 or 1132819-27-2):ti,ab

#31 (CER or CZP):ti,ab

#32 cimzia:ti,ab,kw

#33 (anti-certolizumab* or anticertolizumab* or (anti near/3 certolizumab*)):ti,ab,kw

#34 golimumab:ti,ab,kw

#35 ("cnto 148" or cnto148 or cnto-148 or 476181-74-5):ti,ab

#36 (GOL or GLM):ti,ab

#37 simponi:ti,ab,kw

#38 (anti-golimumab* or antigolimumab* or (anti near/3 golimumab*)):ti,ab,kw

#39 (biologic* near/2 agent*):ti,ab,kw

#40 (CT-P13 or CTP13 or "CT P13" or SB2 or SB-2 or "SB 2" or SB4 or SB-4 or "SB 4" or SB-5 or SB5 or "SB 5"):ti,ab

#41 (biosimilar* or "bio* similar*"):ti,ab,kw

#42 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41

#43 MeSH descriptor: [Enzyme-Linked Immunosorbent Assay] explode all trees

#44 (immundiagnostik* or immunodiagnostik* or immunediagnostik*):ti,ab,kw

#45 "biohit healthcare":ti,ab,kw

#46 (proteomika* or promonitor*):ti,ab,kw

#47 (enzyme* near/3 immunoassay*):ti,ab,kw

#48 (enzyme* near/3 ("immuno* assay*" or "immuno* test*")):ti,ab,kw

#49 ELISA*:ti,ab,kw

#50 (idkmonitor* or (idk near/3 monitor*) or idk-monitor*):ti,ab,kw

#51 ((lisa near/3 tracker*) or lisa-tracker* or lisatracker*):ti,ab,kw

#52 (ridascreen* or (rida near/3 screen*) or rida-screen*):ti,ab,kw

#53 (mabtrack* or (mab near/3 track*) or mab-track*):ti,ab,kw

#54 (sanquin or theradiag):ti,ab,kw

#55 (grifols or progenika):ti,ab,kw

#56 (r-biopharm or rbiopharm or "r biopharm"):ti,ab,kw

#57 ((drug* or trough) near/3 (level* or concentration)):ti,ab,kw

#58 #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57

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

#60 RA:ti,ab

#61 Rheumarthrit*.ti,ab,kw

#62 ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) near/4 (arthrit* or arthros* or polyarthrit* or factor*)):ti,ab,kw

#63 (Chronic* near/4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)):ti,ab,kw

#64 ((Inflammat* or pain* or swell* or stiff*) near/4 (joint* or synovial*)):ti,ab,kw

#65 (Beauvais* near/2 disease*):ti,ab,kw

#66 #59 or #60 or #61 or #62 or #64 or #65

#67 #42 and #58 and #66

Database: Web of Science (SCI and CPCI-S)

Host: Thomson Reuters

Data Parameters: n/a

Date Searched: 30/7/2018

Searcher: SR

Hits: 187

Strategy:

1. #1 TS=(anti-TNF* or antiTNF* or (TNF* near/1 (inhibit* or block*))) OR TS=tumo$r* necrosis* factor* alpha OR TS= (biologic* near/1 DMARD*) OR TS=(biologic* near/3 antirheumati*) OR TS=(anti rheumati* near/3 biologic*) OR TS=(disease modify* near/3 biologic*) OR TS=anti* drug* antibod* OR TS=ADAb OR TS=anti* tumo$r* necrosis* factor* OR TS=monoclonal antibod*
2. #2 TS=etanercept OR TS=(tnr001 or tnr 001 or tnr-001 or 185243-69-0) OR TS=(ETA or ETN) OR TS=(enbrel or erelzi or benepali or lifmior or brenzys) OR TS=(anti-etanercept* or antietanercept* or anti near/2 etanercept*)
3. #3 TS=adalimumab OR TS=(d 2e7 or d2e7 or d-2e7 or 331731-18-1) OR TS=(ADA or ADL or ADM) OR TS=(humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz) OR TS=(anti-adalimumab* or antiadalimumab* or anti near/2 adalimumab*)
4. #4 TS= infliximab OR TS=(170277-31-3 or ta650 or ta 650 or ta-650) OR TS=(INF or IFX) OR TS=(anti-infliximab* or antiinfliximab* or anti near/2 infliximab*) OR TS=(remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi)
5. #5 TS=certolizumab OR TS=(cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2) OR TS=(CER or CZP) OR TS=cimzia OR TS=(anti-certolizumab* or anticertolizumab* or anti near/2 certolizumab*)
6. #6 TS=golimumab OR TS=(cnto 148 or cnto148 or cnto-148 or 476181-74-5) OR TS=(GOL or GLM) OR TS=simponi OR TS=(anti-golimumab* or antigolimumab* or anti near/2 golimumab*)
7. #7 TS=(biologic* near/1 agent*) OR TS=(CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5) OR TS=(biosimilar* or bio* similar*)
8. #8 #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1
9. #9 TS=(immundiagnostik* or immunodiagnostik* or immunediagnostik*) OR TS=biohit healthcare OR TS=(proteomika* or promonitor*) OR TS=(enzyme*

near/2 immunoassay*) OR TS=(enzyme* near/2 immuno* assay*) OR TS=(enzyme* near/2 immuno* test*) OR TS=ELISA*

10. #10 TS= (idkmonitor* or idk near/2 monitor* or idk-monitor*) OR TS=(lisa near/2 tracker* or lisa-tracker* or lisatracker*) OR TS=(ridascreen* or rida near/2 screen* or rida-screen*) OR TS=(mabtrack* or mab near/2 track* or mab-track*) OR TS=(sanquin or theradiag) OR TS=(grifols or progenika) OR TS=(r-biopharm or rbiopharm or r biopharm) OR TS= ((drug* or trough) near/2 (level* or concentration))
11. #11 #10 OR #9
12. #12 TS=RA OR TS=Rheumarthrit* OR TS=((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) near/3 (arthrit* or arthros* or polyarthrit* or factor*)) OR TS=(chronic* near/3 polyarthrit*) OR TS=(chronic* near/3 poly arthrit*) OR TS=(chronic* near/3 rheumati*) OR TS=((Inflammat* or pain* or swell* or stiff*) near/3 (joint* or synovial*)) OR TS=(Beauvais* adj2 disease*)
13. #12 AND #11 AND #8
14. TS=(quality of life OR quality adjusted life OR qaly* OR qald* OR qale* OR qtime* OR life year OR life years OR disability adjusted life OR daly* OR sf36 OR sf 36 OR short form 36 OR shortform 36 OR short form36 OR shortform36 OR sf6 OR sf 6 OR short form 6 OR sf6d OR sf 6d OR short form 6d OR sf8 OR sf 8 OR short form 8 OR sf12 OR sf 12 OR short form 12 OR sf16 OR sf 16 OR sf20 OR sf 20 OR short form 20 OR hql OR hqol OR h qol OR hrqol OR hr qol OR hye OR hyes OR healthy year equivalent* OR healthy years equivalent* OR pqol OR qls OR quality of well being OR index of wellbeing OR qwb OR nottingham health profile* OR sickness impact profile OR health utilit* OR health status OR disutilit* OR rosser OR willingness to pay OR standard gamble* OR time trade off OR time tradeoff OR tto OR hui OR hui1 OR hui2 OR hui3 OR euroqol OR euro qol OR eq5d OR eq 5d OR euroqual OR euro qual OR duke health profile OR functional status questionnaire OR dartmouth coop functional health assessment* OR (utilit* AND (valu* OR measur* OR health OR life OR estimat* OR elicit* OR disease OR score* OR weight)) OR (preference* AND (valu* OR measur* OR health OR life OR estimat* OR elicit* OR disease OR score* OR instrument OR instruments)))
15. #14 and #15

Indexes=SCI-EXPANDED, CPCI-S Timespan=1900-2018;

Website: ScHARR HUD

Date searched: 30/7/2018

Searcher: SR

Hits: 33

Website: HERC Oxford

Date searched: 30/7/2018

Searcher: SR

Hits: 1

Website: EQ-5D EuroQol

Date searched: 30/7/2018

Searcher: SR

Hits: 174

CEA Registry

Date searched: 30/7/2018

Searcher: SR

Hits: 103

Number of hits per database and in total

Database	Hits
MEDLINE	136
MEDLINE In-Process	2
EMBASE	64
Cochrane - NHS EED	0
Web of Science	187
HUD - ScHARR	33
HERC - Oxford	1
EQ-5D - EuroQol	174
CEA Registry	103
Total records	**700**
Duplicates	**70**
Total unique records	**630**

Appendix 2. Included and excluded studies

Table 76: Studies included in the clinical-effectiveness systematic review

	Source	Title	Article type	Contributed data
C. G. Arango, M. L. G. Vivar, E. U. Angulo, I. Gorostiza, C. E. Perez, J. R. De Dios, B. Alvarez, A. R. Escribano, C. Stoye, M. Vasques, J. B. Otano, A. Escobar, Z. Trancho, A. R. Del Agua, L. Del Rio, C. Jorquera, A. Martinez and D. Nagore	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Prospective, intervention, multicenter, non-inferiority study of utility of therapeutic drug monitoring with respect to the efficacy and cost of adalimumab tapering in patients with rheumatic diseases	Conference abstract	Yes
D. Y. Chen, Y. M. Chen, T. Y. Hsieh, W. T. Hung, C. W. Hsieh, H. H. Chen, K. T. Tang and J. L. Lan	Rheumatology	Drug trough levels predict therapeutic responses to dose reduction of adalimumab for rheumatoid arthritis patients during 24 weeks of follow-up	Full article	Yes
I. Gorostiza, E. U. Angulo, C. G. Arango, C. E. Perez, J. R. De Dios, B. Alvarez, A. R. Escribano, C. Stoye, M. Vasques, J. B. Otano, A. Escobar, Z. Trancho, A. R. Del Agua, L. Del Rio, A. Martinez and D. Nagore	Arthritis and Rheumatology	Prospective, intervention, multicenter study of utility of biologic drug monitoring with respect to the efficacy and cost of adalimumab tapering in patients with rheumatic diseases (34-week descriptive data)	Conference abstract	Yes
J. Inciarte-Mundo, M. Hernandez, V. Ruiz-Esquide, J. Ramirez, A. Cuervo, S. Cabrera-Villalba, M. Pascal, J. Yague, J. Canete and R. Sanmarti	Annals of the Rheumatic Diseases	Prediction of flare in rheumatoid arthritis and psoriatic arthritis patients with low disease activity receiving TNF inhibitors: Role of calprotectin and drug trough serum levels. A one-year prospective cohort study	Conference abstract	Yes
M. T. Lopez-Casla, D. Pascual-Salcedo, C. Plasencia, P. Alcozer, S. Garcia-Carazo, G. Bonilla, A. Villalba, D. Peiteado, F. Arribas, E. Martin-Mola and A. Balsa	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	The infliximab dose increase is not correlated with clinical improvement in RA patients	Conference abstract	Yes

	Source	Title	Article type	Contributed data
B. Paredes, C. Plasencia, D. Pascual-Salcedo, I. Monjo, A. Pieren, E. Moral, C. Tornero, P. Bogas, G. Bonilla, L. Nuno, A. Villalba, D. Peiteado, S. Ramiro, T. Jurado, J. Diez, E. Martin-Mola and A. Balsa	Annals of the Rheumatic Diseases	Influence of tapering biological therapies in immunogenicity in a cohort of rheumatoid arthritis with low disease activity	Conference abstract	Yes
B. Paredes, C. Plasencia, D. Pascual-Salcedo, I. Monjo, A. Pieren, E. Moral, C. Tornero, G. Bonilla, L. Nuno, A. Villalba, D. Peiteado, S. Ramiro, T. Jurado, J. Diez, E. Martin-Mola and A. Balsa	Annals of the Rheumatic Diseases	Influence of optimization of biological therapies on immunogenicity in a cohort of rheumatoid arthritis with low disease activity	Conference abstract	Yes
D. Pascual-Salcedo, C. Plasencia, L. Gonzalez Del Valle, T. Lopez Casla, F. Arribas, A. Villalba, G. Bonilla, E. Lopez Granados, E. Martin Mola and A. Balsa	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Therapeutic drug monitoring (TDM) in rheumatic day clinic enables to reduce pharmaceutical cost maintaining clinical efficacy	Conference	Yes
J. Rosas, F. Llinares-Tello, J. Miguel Senabre, G. Santos-Soler, E. Salas-Heredia, X. Barber, A. Pons, C. Cano, M. Lorente and J. Molina	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Economic impact of decreasing adalimumab and etanercept doses and drug monitoring in patients with rheumatoid arthritis in clinical remission: Preliminary study from a local biologics unit	Conference abstract	Yes
J. M. Senabre Gallego, J. Rosas Gomez De Salazar, M. Marco Mingot, A. Naranjo, F. Llinares-Tello, A. Pons, X. Barber-Valles, G. Santos-Soler, E. Salas-Heredia, C. Cano, M. Lorente, J. A. Garcia Gomez and J. Molina	Annals of the Rheumatic Diseases	Clinical activity, ultrasound assessment and drug monitoring in rheumatoid arthritis patients receiving anti-TNF-alpha therapy with extended interval of administration	Conference abstract	Yes

Table 77: Excluded studies (with reasons)

Authors	Source	Title	Reasons for exclusion
R.; Rodriguez-Vidal Alcobendas, A.; Pascual-Salcedo, D.; Murias, S.; Remesal, A.; Diego, C.; Merino, R.	Clinical & Experimental Rheumatology	Monitoring serum etanercept levels in juvenile idiopathic arthritis: a pilot study	Population
P.; Plasencia Alcocer, C.; Pascual, D.; Garcia Carazo, S.; Franco, K. N.; Cagijas, D.; Lojo, L.; Bonilla, G.; Nuno, L.; Villalba, A.; Lopez Casla, M. T.; Balsa, A.; Martin Mola, E.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Imnunogenicity and clinical practice in patients treated with anti-TNF therapy	Population
P.; Plasencia Alcocer, C.; Pascual, D.; Garcia Carazo, S.; Franco, K. N.; Cagijas, D.; Lojo, L.; Bonilla, G.; Nuno, L.; Villalba, A.; Lopez Casla, M. T.; Balsa, A.; Martin Mola, E.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Imnunogenicity and clinical practice in patients treated with anti-TNF therapy	Population
C.; Scrivo Alessandri, R.; Spinelli, F. R.; Ceccarelli, F.; Magrini, L.; Priori, R.; Valesini, G.	Autoimmunity, Part B Novel Applications of Basic Research	Autoantibody production in anti-TNF-alpha-treated patients	Design
A.; Rivera Ametzazurra, N.; Balsa, A.; Arreba, M. P.; Ruiz, E.; Plasencia, C.; Ortiz, J.; Pascual-Salcedo, D.; Munoz, M. C.; De Aysa, C.; Allande, M. J.; Torres, N.; Hernandez, A. M.; Recalde, X.; Martinez, A.; Nagore, D.	Annals of the Rheumatic Diseases	Point-of-care monitoring of anti-infliximab antibodies in patients treated with the reference infliximab or CT-P13 in routine clinical practice	Population
C.; Pomirleanu Ancuta, C.; Belibou, C.; Maxim, R.; Petrariu, L.; Strugariu, G.; Chirieac, R.	Annals of the Rheumatic Diseases	Clinical outcomes of immunogenicity in rheumatoid arthritis patients under anti-TNF biologics: Results from an observational study	Population
C.; Pomirleanu Ancuta, C.; Maxim, R.; Ancuta, E.; Iordache, C.; Dascalu, C.; Chirieac, R.	Revista De Chimie	Clinical Relevance of Rituximab Immunogenicity in Rheumatoid Arthritis A pilot study	Intervention

Authors	Source	Title	Reasons for exclusion
I.; Kapleryte Arstikyte, G.; Butrimiene, I.; Venalis, A.	BioMed Research International	Influence of Immunogenicity on the Efficacy of Long-Term Treatment with TNF alpha Blockers in Rheumatoid Arthritis and Spondyloarthritis Patients	Population
A. S.; Aleksandrova Avdeeva, E. N.; Novikov, A. A.; Karateev, D. E.; Luchihina, E. L.; Cherkasova, M. V.; Nasonov, E. L.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Association of clinical efficacy with serum level of adalimumab (ADA) and anti-adalimumab antibody levels in patients with early rheumatoid arthritis (RA)	Population
W.; Pilari Awni, S.; Ahmed, G.; Noertersheuser, P.	Arthritis and Rheumatism	The effect of methotrexate on adalimumab pharmacokinetics: Pooled analysis of adalimumab pharmacokinetics in patients with rheumatoid arthritis after subcutaneous administration	Design
L. I.; Solberg Bader, S. M.; Kaada, S. H.; Bolstad, N.; Warren, D. J.; Gavasso, S.; Gjesdal, C. G.; Vedeler, C. A.	Scandinavian Journal of Immunology	Assays for Infliximab Drug Levels and Antibodies: AMatter of Scales and Categories	Design
A.; Sanmarti Balsa, R.; Rosas, J.; Castro, S. G.; Cabez, A.; Martin, V.; Montoro, M.	Arthritis and Rheumatology	Immunogenicity of anti-TNF therapies in patients with inflammatory rheumatic diseases and secondary failure: A multicentre study of 570 patients	Outcome
A.; Sanmarti Balsa, R.; Rosas, J.; Martin, V.; Cabez, A.; Gomez, S.; Montoro, M.	Rheumatology	Drug immunogenicity in patients with inflammatory arthritis and secondary failure to tumour necrosis factor inhibitor therapies: the REASON study	Outcome
S.; Salvatierra Bandres Ciga, J.; Lopez-Sidro, M.; Garcia-Sanchez, A.; Duran, R.; Vives, F.; Raya-Alvarez, E.	JCR: Journal of Clinical Rheumatology	An examination of the mechanisms involved in secondary clinical failure to adalimumab or etanercept in inflammatory arthropathies	Design

Authors	Source	Title	Reasons for exclusion
S.; Salvatierra Ossorio Bandres Ciga, J.; Lopez-Sidro, M.; Garcia Sanchez, A.; Duran Ogalla, R.; Vives Montero, F.; Raya-Alvarez, E.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	The utility of the mechanistic model in inflammatory arthropaties with secondary clinical failure to adalimumab, but not to etanercept	Design
F. I.; Krauchi Bantleon, S.; Schuster, T. B.; Schneider, M.; Abel Buhlmann, A.	Journal of Crohn's and Colitis	Quantum blue adalimumab: Development of the first point of care rapid test for therapeutic drug monitoring of serum adalimumab levels	Design
F. I.; Krauchi Bantleon, S.; Schuster, T. B.; Schneider, M.; Weber, J. M.	Annals of the Rheumatic Diseases	Quantum blue adalimumab: Evaluation of a point of care rapid test for therapeutic drug monitoring of serum adalimumab levels	Design
S.; Plasencia Baos, C.; Ramiro, S.; Moral, R.; Diez, J.; Martin-Mola, E.; Balsa, A.	Arthritis and Rheumatism. Conference: Annual Scientific Meeting of the American College of Rheumatology and Association of Rheumatology Health Professionals	Effect on rheumatoid factor and anti-cyclic citrullinated peptide antibodies levels of treatment with infliximab and adalimumab in patients with rheumatoid arthritis	Design
N. L.; Mohammed Barlow, P.; Berg, J. D.	Clinical Chemistry and Laboratory Medicine	Clinical study of serum trough infliximab concentrations and anti-infliximab antibodies in a cohort of gastroenterology and rheumatology patients	Design
N. L.; Mohammed Barlow, P.; Berg, J. D.	Annals of Clinical Biochemistry	Serum trough infliximab and anti-infliximab antibodies in a cohort of gastroenterology and rheumatology patients' infliximab therapeutic drug monitoring	Design
G. M.; de Groot Bartelds, E.; Nurmohamed, M. T.; Hart, M. H.; van Eede, P. H.; Wijbrandts, C. A.; Crusius, J. B.; Dijkmans, B. A.; Tak, P. P.; Aarden, L.; Wolbink, G. J.	Arthritis Research & Therapy	Surprising negative association between IgG1 allotype disparity and anti-adalimumab formation: a cohort study	Design
G. M.; Krieckaert Bartelds, C. L.; Nurmohamed, M. T.; Van	Arthritis and Rheumatism	Immunogenicity in a 3-year follow-up cohort of adalimumab treated rheumatoid arthritis patients	Design

Authors	Source	Title	Reasons for exclusion
Schouwenburg, P.; Dijkmans, B. A.; Wolbink, G. J.			
G. M.; Krieckaert Bartelds, C. L.; Nurmohamed, M. T.; van Schouwenburg, P. A.; Lems, W. F.; Twisk, J. W.; Dijkmans, B. A.; Aarden, L.; Wolbink, G. J.	JAMA	Development of antidrug antibodies against adalimumab and association with disease activity and treatment failure during long-term follow-up	Population
G. M.; Wolbink Bartelds, G. J.; Stapel, S.; Aarden, L.; Lems, W. F.; Dijkmans, B. A. C.; Nurmohamed, M. T.	Annals of the Rheumatic Diseases	High levels of human anti-human antibodies to adalimumab in a patient not responding to adalimumab treatment [4]	Design
C.; Ruiz Bastida, V.; Pascal, M.; Yague, J.; Sanmarti, R.; Soy, D.	British Journal of Clinical Pharmacology	Is there potential for therapeutic drug monitoring of biologic agents in rheumatoid arthritis?	Design
N. K.; Heilig Bender, C. E.; Droll, B.; Wohlgemuth, J.; Armbruster, F. P.; Heilig, B.	Rheumatology International	Immunogenicity, efficacy and adverse events of adalimumab in RA patients	Design
K. Bendtzen	Arthritis and Rheumatism	Is there a need for immunopharmacologic guidance of anti-tumor necrosis factor therapies?	Design
K. Bendtzen	Immunotherapy	Anti-TNF-alpha biotherapies: Perspectives for evidence-based personalized medicine	Design
K. Bendtzen	Discovery Medicine	Personalized Medicine: Theranostics (Therapeutics Diagnostics) Essential for Rational Use of Tumor Necrosis Factor-alpha Antagonists	Design
M.; Damiani Benucci, A.; Li Gobbi, F.; Bandinelli, F.; Infantino, M.; Grossi, V.; Manfredi, M.; Noguier, G.; Meacci, F.	Biologics	Correlation between HLA haplotypes and the development of antidrug antibodies in a cohort of patients with rheumatic diseases	Design

Authors	Source	Title	Reasons for exclusion
M.; Gobbi Benucci, F. L.; Meacci, F.; Manfredi, M.; Infantino, M.; Severino, M.; Testi, S.; Sarzi-Puttini, P.; Ricci, C.; Atzeni, F.	Biologics: Targets and Therapy	Antidrug antibodies against TNF-blocking agents: Correlations between disease activity, hypersensitivity reactions, and different classes of immunoglobulins	Design
M.; Infantino Benucci, M.; Manfredi, M.; Olivito, B.; Sarzi-Puttini, P.; Atzeni, F.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Anti-drug-antibodies but not IGG-4 antibodies against TNF blockers influence the activity of anti-TNF drugs in rheumatoid arthritis	Design
M.; Li Gobbi Benucci, F.; Meacci, F.; Manfredi, M.; Infantino, M.; Severino, M.; Testi, S.; Sarzi-Puttini, P.; Ricci, C.; Atzeni, F.	Biologics	Antidrug antibodies against TNF-blocking agents: correlations between disease activity, hypersensitivity reactions, and different classes of immunoglobulins	Design
E.; Mansson Berthold, B.; Gullstrand, B.; Geborek, P.; Saxne, T.; Bengtsson, A. A.; Kahn, R.	Scandinavian Journal of Rheumatology	Tumour necrosis factor-alpha/etanercept complexes in serum predict long-term efficacy of etanercept treatment in seronegative rheumatoid arthritis	Design
C. O.; Ince Bingham, A.; Haraoui, B.; Keystone, E. C.; Chon, Y.; Baumgartner, S.	Current Medical Research and Opinion	Effectiveness and safety of etanercept in subjects with RA who have failed infliximab therapy: 16-week, open-label, observational study	Design
P.; Plasencia Bogas, C.; Pascual-Salcedo, D.; Bonilla, G.; Moral, E.; Tornero, C.; Nuno, L.; Villalba, A.; Peiteado, D.; Martinez, A.; Hernandez, B.; Balsa, A.	Annals of the Rheumatic Diseases	Discontinuation of first biologic therapy in rheumatoid arthritis: Main causes and correlation between secondary inefficacy and development of immunogenicity	Design
P.; Plasencia Bogas, C.; Pascual-Salcedo, D.; Bonilla, G.; Moral, E.; Tornero, C.; Nuno, L.; Villalba, A.; Peiteado, D.; Martinez, A.; Hernandez, B.; Balsa, A.	Annals of the Rheumatic Diseases	Influence of immunogenicity to the first TNF-I therapy on response to the second biologic agent in RA patients	Design

Authors	Source	Title	Reasons for exclusion
P.; Plasencia-Rodriguez Bogas, C.; Balsa, A.; Pascual-Salcedo, D.; Bonilla, G.; Coro, E. M.; Tornero, C.; Nuno, L.; Peiteado, D.; Martinez, A.; Hernandez, B.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Influence of immunogenicity to the first anti-TNF therapy on response to the second biologic agent in RA patients	Design
Y.; Ben Horin Braun-Moscovici, S.; Dagan, A.; Toledano, K.; Markovits, D.; Saffouri, A.; Beshara, R.; Rozin, A.; Nahir, M. A.; Chowers, Y.; Balbir-Gurman, A.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	The input of measuring of infliximab and adalimumab levels and levels of antibodies to these drugs in the management of patients with autoimmune diseases treated with anti TNF monoclonal antibodies	Design
F.; Cao Cao, H. L.; Cao, X. C.	International Journal of Clinical Pharmacology and Therapeutics	A review of six methods for monitoring infliximab concentrations and antibodies to infliximab	Design
M.; Ramsey Casal, M.; Moreland, L. W.; Fernandez, C.	Arthritis and Rheumatology	A cytometric assay for monitoring adalimumab immunogenicity and drug concentrations can distinguish anti-adalimumab antibodies from interference	Design
N. V.; Buurman Casteele, D. J.; Sturkenboom, M. G. G.; Kleibeuker, J. H.; Vermeire, S.; Rispens, T.; van der Kleij, D.; Gils, A.; Dijkstra, G.	Alimentary Pharmacology & Therapeutics	Detection of infliximab levels and anti-infliximab antibodies: a comparison of three different assays	Design
M. J. Cates	Rheumatology (United Kingdom)	Anti-tumour necrosis factor a drug levels and anti-drug antibodies in guiding clinical decision making in rheumatology: A draft algorithm and illustrative cases	Design
S.; Payne Cavan, K.; Barton, A.	Value in Health	MEASURING ADALIMUMAB DRUG LEVELS BY ELISA TO DETECT TREATMENT RESPONSE IN RHEUMATOID ARTHRITIS: A SYSTEMATIC REVIEW AND BIVARIATE META-ANALYSIS	Design

Authors	Source	Title	Reasons for exclusion
P. R.; Pascual-Salcedo Chamaida, D.; Bonilla, M.; Villalba, A.; Lopez-Casla, M.; Peiteado, D.; Garcia-Carazo, S.; Ramiro, S.; Franco, K.; Cajigas, D.; Martin-Mola, E.; Balsa, A.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	The early infliximab levels monitoring can predict the developement of anti-drug antibodies in a cohort of rheumatoid arthritis patients treated with infliximab	Design
E.; Mulleman Chasseuil, D.; Aubourg, A.; Lecomte, T.; Paintaud, G.; Ternant, D.	Fundamental and Clinical Pharmacology	Determination of infliximab cut-off concentrations predicting presence or absence of antibodies towards infliximab (ATI) in chronic inflammatory diseases	Design
K. Chatzidionysiou	Scandinavian Journal of Rheumatology	Optimizing biological treatments for rheumatoid arthritis	Design
D. Y.; Chen Chen, Y. M.; Tsai, W. C.; Tseng, J. C.; Chen, Y. H.; Hsieh, C. W.; Hung, W. T.; Lan, J. L.	Annals of the Rheumatic Diseases.	Significant associations of antidrug antibody levels with serum drug trough levels and therapeutic response of adalimumab and etanercept treatment in rheumatoid arthritis	Population
D. Y.; Chen Chen, Y. M.; Hung, W. T.; Chen, H. H.; Hsieh, C. W.; Chen, Y. H.; Huang, W. N.; Hsieh, T. Y.	Annals of the Rheumatic Diseases	Immunogenicity, drug trough levels and therapeutic response in patients with rheumatoid arthritis or ankylosing spondylitis after 24-week golimumab treatment	Population
C. B.; Favalli Chighizola, E. G.; Meroni, P. L.	Clinical Reviews in Allergy & Immunology	Novel mechanisms of action of the biologicals in rheumatic diseases	Design
S.; Nicaise-Roland Chollet-Martin, P.; De Chaisemartin, L.; Grootenboer-Mignot, S.; Hayem, G.; Pelletier, A. L.; Amiot, A.; Descamps, V.; Bouhnik, Y.; Meyer, O.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Simultaneous determination of anti-infliximab antibodies and residual infliximab levels to monitor anti-TNF therapy	Outcome
V.; Kaliyaperumal Chow, A.; Zhang, N.; Miller, J.; Mytych, D.; Starcevic Manning, M.; Wala, I.; Wang, H.; Krishnan, E.	Journal of Crohn's and Colitis	Development of anti-drug antibodies among those treated with adalimumab and ABP 501 and its impact on serum drug concentration in randomised controlled studies	Design

Authors	Source	Title	Reasons for exclusion
S. B.; Salvatierra Ciga, J.; Lopez-Sidro, M.; Garcia-Sanchez, A.; Duran, R.; Vives, F.; Raya-Alvarez, E.	Jcr-Journal of Clinical Rheumatology	An Examination of the Mechanisms Involved in Secondary Clinical Failure to Adalimumab or Etanercept in Inflammatory Arthropathies	Design
E. W. S.; Wagner Clair, C. L.; Fasanmade, A. A.; Wang, B.; Schaible, T.; Kavanaugh, A.; Keystone, E. C.	Arthritis and Rheumatism	The relationship of serum infliximab concentrations to clinical improvement in rheumatoid arthritis - Results from ATTRACT, a multicenter, randomized, double-blind, placebo-controlled trial	Population
I.; Spinelli Cludts, F. R.; Morello, F.; Hockley, J.; Valesini, G.; Wadhwa, M.	Cytokine	Anti-therapeutic antibodies and their clinical impact in patients treated with the TNF antagonist adalimumab.[Reprint in Cytokine. 2018 Jan;101:70-77; PMID: 29174881]	Population
I.; Spinelli Cludts, F. R.; Morello, F.; Hockley, J.; Valesini, G.; Wadhwa, M.	Cytokine	Reprint of "Anti-therapeutic antibodies and their clinical impact in patients treated with the TNF antagonist adalimumab".[Reprint of Cytokine. 2017 Aug;96:16-23; PMID: 28279855]	Population
J.; Couble Collet-Brose, P. J.; Deehan, M. R.; Nelson, R. J.; Ferlin, W. G.; Lory, S.	Journal of Immunology Research	Evaluation of Multiple Immunoassay Technology Platforms to Select the Anti-Drug Antibody Assay Exhibiting the Most Appropriate Drug and Target Tolerance	Design
F.; Cetin Cosan, E. A.; Gazioglu, S. B.; Yazici, A.; Yilmazer, B.; Cefle, A.; Deniz, G.	Clinical and Experimental Rheumatology	How could be used the autoantibodies against anti-TNF agents in clinical practice? Two years follow-up study	Population
C. I.; Daien Daien, V.; Parussini, E.; Dupuy, A. M.; Combe, B.; Morel, J.	Journal of Rheumatology	Etanercept concentration in patients with rheumatoid arthritis and its potential influence on treatment decisions: a pilot study	Design
C. W. N.; Schellens Damen, J. H. M.; Beijnen, J. H.	Human Antibodies	Bioanalytical methods for the quantification of therapeutic monoclonal antibodies and their application in clinical pharmacokinetic studies	Design

Authors	Source	Title	Reasons for exclusion
F.; Bian Darrouzain, S. M.; Desvignes, C.; Bris, C.; Watier, H.; Paintaud, G.; de Vries, A.	Therapeutic Drug Monitoring	Immunoassays for Measuring Serum Concentrations of Monoclonal Antibodies and Anti-biopharmaceutical Antibodies in Patients	Design
Jmgr; Pascual-Salcedo de Morales, D.; Tello, F. L.; Mendez, L. V.	Medicina Clinica	Anti-tumor necrosis factor drug therapy: The usefulness of monitoring drug levels and anti-drug antibodies in clinical practice	Design
J. R. G.; Llinares-Tello De Salazar, F.; Senabre-Gallego, J. M.; Santos-Soler, G.; Santos-Ramirez, C.; Salas-Heredia, E.; Barber-Valles, X.	Arthritis and Rheumatism. Conference: Annual Scientific Meeting of the American College of Rheumatology and Association of Rheumatology Health Professionals	Evaluation of anti-tumor necrosis factor levels and anti-tumor necrosis factor antibodies in rheumatic diseases treated with infliximab and adalimumab; preliminary results from a local registry	Population
A. A.; van Herwaarden den Broeder, N.; van den Bemt, B. J. F.	Current Opinion in Rheumatology	Therapeutic drug monitoring of biologicals in rheumatoid arthritis: a disconnect between beliefs and facts	Design
D.; Rinaudo Denarie, M.; Thomas, T.; Paul, S.; Marotte, H.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Longitudinal study of serum TNF alpha levels, infliximab, and antibodies to infliximab in rheumatoid arthritis	Population
D.; Rinaudo Denarie, M.; Thomas, T.; Paul, S.; Marotte, H.	Annals of the Rheumatic Diseases	Methotrexate reduced TNF bioactivity by anti-infliximab antibody prevention in rheumatoid arthritis patients treated with infliximab	Design
D.; Rinaudo-Gaujous Denarie, M.; Thomas, T.; Paul, S.; Marotte, H.	Mediators of Inflammation	Methotrexate Reduced TNF Bioactivity in Rheumatoid Arthritis Patients Treated with Infliximab	Design
T.; Weinblatt Dervieux, M. E.; Kivitz, A.; Kremer, J. M.	Annals of the Rheumatic Diseases	Methotrexate polyglutamation in relation to infliximab pharmacokinetics in rheumatoid arthritis	Design
M.; Iliuta Diana, M.; Gainaru, C.; Apetrei, N.; Luca, G.; Groseanu, L.; Saulescu, I.; Constantinescu, C.; Bojinca, V.; Borangiu, A.; Balanescu,	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Clinical utility of measuring drug and anti-drug antibody concentration of biologic agents in	Design

Authors	Source	Title	Reasons for exclusion
A.; Predeteanu, D.; Ionescu, R.; Opris, D.		rheumatoid arthritis patients with moderate and high disease activity	
F.; Ataman Doghanji, S.; Ozdemirel, A. E.; Seckin, R. B.; Yalcin, A. P.; Bavbek, S.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Relationship between immunogenicity, hypersensitivity reactions and skin tests against infliximab, etanercept and adalimumab in patients with rheumatoid arthritis and ankylosing spondylitis	Design
S.; Beuermann Drynda, R.; Kekow, J.	Annals of the Rheumatic Diseases	Determination of anti-drug antibodies in long-term treatment of rheumatoid arthritis patients with etanercept	Design
S.; Kekow Drynda, J.	Zeitschrift Fur Rheumatologie	Determination of TNF alpha blocker serum levels and anti drug antibodies during long term treatment of rheumatoid arthritis patients and their association with clinical outcome and selected biomarkers	Design
S.; Kekow Drynda, J.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Clinical relevance of etanercept levels and anti-etanercept antibodies in long-term treatment of rheumatoid arthritis patients	Design
S.; Kekow Drynda, J.	Annals of the Rheumatic Diseases	Clinical importance of anti-drug and serum drug level testing in rheumatoid arthritis patients treated with etanercept	Population
E.; Mulleman Ducourau, D.; Paintaud, G.; Lin, D. C. M.; Lauferon, F.; Ternant, D.; Watier, H.; Goupille, P.	Arthritis Research & Therapy	Antibodies toward infliximab are associated with low infliximab concentration at treatment initiation and poor infliximab maintenance in rheumatic diseases	Population
E.; Ternant Ducourau, D.; Corondan, A.; Legoff, B.; Perdriger, A.; Devauchelle, V.; Solau-Gervais, E.	Arthritis and Rheumatism	Body surface area, erythrocyte sedimentation rate, methotrexate and antibodies to infliximab influence the pharmacokinetics of infliximab in rheumatoid arthritis	Population

Authors	Source	Title	Reasons for exclusion
E.; Ternant Ducourau, D.; Mulleman, D.; Mammou, S.; Lin, D. C. M.; Watier, H.; Paintaud, G.	Arthritis and Rheumatism	Antibodies towards infliximab are associated with poor infliximab maintenance and low infliximab concentrations	Design
C.; Dejaco Duftner, C.; Kullich, W.; Klauser, A.; Goldberger, C.; Falkenbach, A.; Schirmer, M.	Annals of the Rheumatic Diseases	Preferential type 1 chemokine receptors and cytokine production of CD28(-) T cells in ankylosing spondylitis	Design
A. F.; Misra Edrees, S. N.; Abdou, N. I.	Clinical & Experimental Rheumatology	Anti-tumor necrosis factor (TNF) therapy in rheumatoid arthritis: correlation of TNF-alpha serum level with clinical response and benefit from changing dose or frequency of infliximab infusions	Population
P.; Burmester Emery, G. R.; Naredo, E.; Zhou, Y.; Hojnik, M.; Conaghan, P. G.	BMJ Open	Design of a phase IV randomised, double-blind, placebo-controlled trial assessing the ImPact of Residual Inflammation Detected via Imaging TEchniques, Drug Levels and Patient Characteristics on the Outcome of Dose TaperIng of Adalimumab in Clinical Remission Rheumatoid ArThritis (RA) patients (PREDICTRA)	Design
N.; De Carvalho Emi Aikawa, J. F.; Artur Almeida Silva, C.; Bonfa, E.	Clinical Reviews in Allergy and Immunology	Immunogenicity of anti-TNF-alpha agents in autoimmune diseases	Design
G. P. Eng	Danish Medical Journal	Optimizing biological treatment in rheumatoid arthritis with the aid of therapeutic drug monitoring	Design
G. P.; Bouchelouche Eng, P.; Bartels, E. M.; Bliddal, H.; Bendtzen, K.; Stoltenberg, M.	PLoS ONE [Electronic Resource]	Anti-Drug Antibodies, Drug Levels, Interleukin-6 and Soluble TNF Receptors in Rheumatoid Arthritis Patients during the First 6 Months of Treatment with Adalimumab or Infliximab: A Descriptive Cohort Study	Design

Authors	Source	Title	Reasons for exclusion
C.; Lind Eriksson, P.; Nystrand, M.; Moverare, R.	Allergy: European Journal of Allergy and Clinical Immunology	A new automated anti-drug antibody screening assay with high sensitivity and drug tolerance	Design
M.; Pistis Fabris, C.; Zabotti, A.; Picco, L.; Curcio, F.; Tonutti, E.; De Vita, S.	Drug Metabolism Letters	The Detection of Anti-adalimumab Antibodies in a Series of Inflammatory Polyarthritis: three ELISA Methods Compared	Design
A. Fogdell-Hahn	Scandinavian Journal of Immunology	Antidrug Antibodies: B Cell Immunity Against Therapy	Design
R.; Shakhnovich Funk, V.; Van Haandel, L.; Becker, M. L.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Infliximab use in JIA and uveitis: Does methotrexate help or hinder?	Population
D. E.; Wallis Furst, R.; Broder, M.; Beenhouwer, D. O.	Seminars in Arthritis & Rheumatism	Tumor necrosis factor antagonists: different kinetics and/or mechanisms of action may explain differences in the risk for developing granulomatous infection	Design
C.; Diana Gainaru, M.; Iliuta, M.; Luca, G.; Apetrei, N.; Constantinescu, C.; Groseanu, L.; Bojinca, V.; Saulescu, I.; Borangiu, A.; Balanescu, A.; Predeteanu, D.; Ionescu, R.; Opris, D.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Infliximab vs etanercept: The importance of immunogenicity and serum drug monitoring in clinical practice	Population
S.; Antunes Garces, M.; Benito-Garcia, E.; Canas-Silva, J.; Aarden, L.; Demengeot, J.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	A preliminary algorithm introducing immunogenicity assessment in the management of RA patients receiving biotechnological therapies	Design
S.; Canas-da-Silva Garces, J.; Aarden, L.; Demengeot, J.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	New algorithm to approach ra patients receiving biologic therapies: Introducing immunogenicity assessment in the eular guidelines	Design

Authors	Source	Title	Reasons for exclusion
S.; Demengeot Garces, J.; Da Silva, J. C.; Aarden, L.	Arthritis and Rheumatism. Conference: Annual Scientific Meeting of the American College of Rheumatology and Association of Rheumatology Health Professionals	Bridging elisa as a screening assay to monitor immunogenicity in routine clinical practice	Design
S.; Demengeot Garces, J.; Wolbink, G. J.; Aarden, L.; Benito-Garcia, E.	Arthritis and Rheumatism. Conference: Annual Scientific Meeting of the American College of Rheumatology and Association of Rheumatology Health Professionals	The immunogenicity of infliximab, adalimumab and etanercept in rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis, crohn's disease and ulcerative colitisa quantitative and a qualitative review	Design
S.; Demengeot Garces, J.; Benito-Garcia, E.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Clinical impact of immunogenicity of infliximab, adalimumab and etanercept: A systematic review of the literature with a meta-analysis	Design
S.; Demengeot Garces, J.; Canas-da-Silva, J.; Aarden, L.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Bridging ELISA as a secreening assay to monitor immunogenicity in routine clinical practice	Design
S.; Demengeot Garces, J.; Benito-Garcia, E.	Annals of the Rheumatic Diseases	The immunogenicity of anti-TNF therapy in immune-mediated inflammatory diseases: a systematic review of the literature with a meta-analysis	Design
S.; Freitas Garces, J.; Canas-Silva, J.; Aarden, L.; Demengeot, J.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	The impact of immunogenicity on drug safety profile	Design
S.; Payne Gavan, K.; Barton, A.	Annals of the Rheumatic Diseases	A systematic review and bivariate meta-analysis of studies that measured adalimumab drug levels by elisa to detect treatment response in rheumatoid arthritis	Design
M. C.; Ogata Genovese, A.; Nomura, A.; Bao, M.; Hitraya, E.; Lacey, S.; Burmester, G.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of	Immunogenicity of subcutaneous and intravenous tocilizumab as monotherapy or in combination with DMARDS	Intervention

Authors	Source	Title	Reasons for exclusion
	Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP		
M.; Iniesta Navalon Gil Candel, C.; Onteniente Candela, M.; Rentero Redondo, L.; Caballero Requejo, C.; Salar Valverde, N.; Gallego Munoz, C.	European Journal of Hospital Pharmacy	Study of the prevalence of immunogenicity in patients treated with anti-tumour necrosis factor monoclonal antibodies	Population
D. D. Gladman	Arthritis & Rheumatology	Clinical Utility of Random Anti-Tumor Necrosis Factor Drug-Level Testing and Measurement of Antidrug Antibodies on the Long-Term Treatment Response in Rheumatoid Arthritis (vol 67, pg 2011, 2015)	Design
B.; Kringelbach Glintborg, T.; Hogdall, E.; Sorensen, I. J.; Jensen, D. V.; Loft, A. G.; Hendricks, O.; Jensen Hansen, I. M.; Bolstad, N.; Gron, K.; Eng, G.; Enevold, C.; Nielsen, C. H.; Warren, D.; Goll, G.; Gehin, J.; Johansen, J. S.; Hetland, M. L.	Annals of the Rheumatic Diseases	Non-medical switch from originator to biosimilar infliximab among patients with inflammatory rheumatic disease-impact on s-infliximab and antidrug-antibodies. Results from the national danish rheumatologic biobank and the danbio registry	Design
G. L.; Jorgensen Goll, K. K.; Sexton, J.; Olsen, I. C.; Bolstad, N.; Lorentzen, M.; Haavardsholm, E. A.; Mork, C.; Jahnsen, J.; Kvien, T. K.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Long-term safety and efficacy of biosimilar infliximab (CT-P13) after switching from originator infliximab: Results from the 26-week open label extension of a randomized Norwegian trial	Design
G. L.; Olsen Goll, I. C.; Jorgensen, K. K.; Lorentzen, M.; Bolstad, N.; Haavardsholm, E. A.; Lundin, K. E. A.; Mork, C.; Jahnsen, J.; Kvien, T. K.	Arthritis and Rheumatology	Biosimilar infliximab (CT-P13) is not inferior to originator infliximab: Results from a 52-week randomized switch trial in Norway	Design
G. L.; Olsen Goll, I. C.; Bolstad, N.; Jorgensen, K. K.; Lorentzen, M.; Mork,	Annals of the Rheumatic Diseases	Disease worsening and safety in patients switching from originator infliximab to biosimilar infliximab (CT-P13) in the nor-	Design

Authors	Source	Title	Reasons for exclusion
C.; Jahnsen, J.; Haavardsholm, E. A.; Kvien, T. K.		switch study: Explorative analysis of RA patients	
G. L.; Olsen Goll, I. C.; Lundin, K. E. A.; Jorgensen, K. K.; Lorentzen, M.; Klaasen, R. A.; Warren, D. J.; Mork, C.; Jahnsen, J.; Haavardsholm, E. A.; Kvien, T. K.; Bolstad, N.	Annals of the Rheumatic Diseases	Immunogenicity in patients switching from stable originator infliximab treatment to CT-P13: Analyses across six diseases from the 52-week randomized nor-switch study	Design
P. A.; Vegh Golovics, Z.; Rutka, M.; Gecse, K.; Balint, A.; Farkas, K.; Banai, J.; Bene, L.; Gasztonyi, B.; Kristof, T.; Lakatos, L.; Miheller, P.; Palatka, K.; Patai, A.; Salamon, A.; Szamosi, T.; Szepes, Z.; Toth, G. T.; Vincze, A.; Biro, E.; Lovasz, B.; Kurti, Z.; Nagy, F.; Molnar, T.; Lakatos, P.	Journal of Crohn's and Colitis	Predicting short and medium-term efficacy of the biosimilar infliximab: Trough levels/do anti-drug antibody's or clinical/biochemical markers play a more important role?	Design
B.; Baltrukonis Gorovits, D. J.; Bhattacharya, I.; Birchler, M. A.; Finco, D.; Sikkema, D.; Vincent, M. S.; Lula, S.; Marshall, L.; Hickling, T. P.	Clinical and Experimental Immunology	Immunoassay methods used in clinical studies for the detection of anti-drug antibodies to adalimumab and infliximab	Design
S.; Bliddal Gudbrandsdottir, H.; Petri, A.; Terslev, L.; Danneskiold-Samsoe, B.; Bjornhart, B.; Bendtzen, K.; Muller, K.	Scandinavian Journal of Rheumatology	Plasma TNF binding capacity profiles during treatment with etanercept in rheumatoid arthritis	Design
M.; Dit Jeanfavre Guirgis, M. F.; Benaim, C.; Perreau, M.; Michetti, P.; Maillard, M.; Zufferey, P.	Arthritis and Rheumatology	Comparison of infliximab immunogenicity in inflammatory arthritis versus inflammatory bowel disease patients in routine clinical practice	Design
H. B.; Bolstad Hammer, N.; Warren, D. J.; Goll, G. L.	Annals of the Rheumatic Diseases	Patients with low serum adalimumab concentrations display poor ultrasonographic	Design

Authors	Source	Title	Reasons for exclusion
		response to treatment; results of a follow-up study of patients with rheumatoid arthritis	
B.; Cameron Haraoui, L.; Ouellet, M.; White, B.	Journal of Rheumatology	Anti-infliximab antibodies in patients with rheumatoid arthritis who require higher doses of infliximab to achieve or maintain a clinical response	Design
M. H.; de Vrieze Hart, H.; Wouters, D.; Wolbink, G. J.; Killestein, J.; de Groot, E. R.; Aarden, L. A.; Rispens, T.	Journal of Immunological Methods	Differential effect of drug interference in immunogenicity assays	Population
S.; Suzuki Hayashi, K.; Yoshimoto, K.; Takeshita, M.; Kurasawa, T.; Yamaoka, K.; Takeuchi, T.	Rheumatology & Therapy	Early Prognostic Factors Associated with the Efficacy of Infliximab Treatment for Patients with Rheumatoid Arthritis with Inadequate Response to Methotrexate	Design
D.; Valor Hernandez, L.; De La Torre, I.; Martinez, L.; Nieto, J. C.; Del Rio, T.; Naredo, E.; Gonzalez, C.; Lopez-Longo, J.; Montoro, M.; Monteagudo, I.; Carreno, L.	Annals of the Rheumatic Disease. Conference: European Workshop for Rheumatology Research	Establishing cut-off of infliximab levels and anti-infliximab antibodies by commercial elisa in patients with rheumatoid arthritis	Design
M.; Boso Herold, L.; Haueis, T.; Klotz, W.; Zangerl, G.	Annals of the Rheumatic Diseases	No need to detect anti-drug antibodies in patients treated with TNF inhibitors	Design
M. L. Hetland	Danish Medical Bulletin	Modern treatment strategies in rheumatoid arthritis: Impact on, and predictors of, disease activity and disease course	Design
D.; Valtanen Ho, S.; Havana, M.; Kroger, L.; Eklund, K.; Jokiranta, S.	Annals of the Rheumatic Diseases	Real-life infliximab and adalimumab trough level and anti-drug antibody measurements in rheumatology: The finnish experience	Design
B. D.; Stamp Hock, L. K.; Hayman, M. W.; Keating, P. E.; Helms, E. T.; Barclay, M. L.	Therapeutic Drug Monitoring	Development of an ELISA-Based Competitive Binding Assay for the Analysis of Drug	Design

Authors	Source	Title	Reasons for exclusion
		Concentration and Antidrug Antibody Levels in Patients Receiving Adalimumab or Infliximab	
B.; O'Donnell J Hock, L.; Liu, J.; Keating, P.; Spellerberg, M.; Stamp, L.; Barclay, M.	Annals of the Rheumatic Diseases	Anti-drug antibodies: Assay performance in patients treated with anti-TNF biodrugs	Design
C.; Brock Hornshoj-Sorensen, B.; Tarp, U.; Pfeiffer-Jensen, M.	Annals of the Rheumatic Diseases	The time window to determine trough values of etanercept is important in personalized medicine regime independently of methotrexate coadministration	Design
M.; Yoshio Hoshino, T.; Onishi, S.; Minota, S.	Modern Rheumatology	Influence of antibodies against infliximab and etanercept on the treatment effectiveness of these agents in Japanese patients with rheumatoid arthritis	Design
A.; Calligaro Hoxha, A.; Tonello, M.; Carletto, A.; Paolazzi, G.; Bortolotti, R.; Felicetti, M.; Ramonda, R.; Del Ross, T.; Grava, C.; Boaretto, M.; Favaro, M.; Teghil, V.; Ruffatti, A.; Punzi, L.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Clinical significance of anti-adalimumab antibodies in rheumatoid arthritis, ankylosing spondilitis and psoriasic arthritis	Design
J.; Hernandez Inciarte-Mundo, M. V.; Cabrera, S.; Ruiz-Esquide, V.; Ramirez, J.; Canete, J. D.; Yague, J.; Sanmarti, R.	Arthritis and Rheumatism	Immunogenicity induced by tumor necrosis factor antagonists in chronic inflammatory arthropathies: Retrospective study in clinical practice conditions	Design
J.; Hernandez Inciarte-Mundo, M. V.; Cabrera-Villalba, S.; Ramirez, J.; Cuervo, A.; Ruiz-Esquide, V.; Gonzalez Navarro, A.; Yague, J.; Canete, J. D.; Sanmarti, R.	Arthritis and Rheumatology	Calprotectin serum levels reflect residual inflammatory activity in patients with rheumatoid arthritis and psoriatic arthritis on clinical remission or low disease activity undergoing TNF-antagonists therapy	Design
J.; Ramirez Inciarte-Mundo, J.; Ruiz-Esquide, V.; Hernandez, M. V.;	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of	Calprotectin and TNF antagonist serum trough levels identify active ultrasound synovitis in	Design

Authors	Source	Title	Reasons for exclusion
Camacho, O.; Cabrera-Villalba, S.; Cuervo, A.; Pascal, M.; Yague, J.; Canete, J. D.; Sanmarti, R.	Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	rheumatoid arhritis and psoriatic arthritis patients in remission or low disease activity	
J.; Ramirez Inciarte-Mundo, J.; Hernandez, M. V.; Ruiz-Esquide, V.; Cuervo, A.; Cabrera-Villalba, S. R.; Pascal, M.; Yague, J.; Canete, J. D.; Sanmarti, R.	Arthritis Research & Therapy	Calprotectin and TNF trough serum levels identify power Doppler ultrasound synovitis in rheumatoid arthritis and psoriatic arthritis patients in remission or with low disease activity	Design
J.; Ramirez Garcia Inciarte-Mundo, J.; Estrada Alarcon, P.; Garcia Manrique, M.; Gonzalez Navarro, A.; Saura, C.; Narvaez, J.; Rodriguez-Moreno, J.; Gomez-Centeno, A.; Yague, J.; Canete, J.; Sanmarti, R.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Drug serum levels of TNF antagonists do not correlate with subclinical synovitis by ultrasound in patients with rheumatoid arthritis and psoriatic arthritis in clinical remission or low disease activity	Design
Y.; Fujii Ishikawa, T.; Kondoh-Ishikawa, S.; Hashimoto, M.; Furu, M.; Ito, H.; Imura, Y.; Nakashima, R.; Yukawa, N.; Yoshifuji, H.; Ohmura, K.; Mimori, T.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Immunogenicity is associated with lupus-like autoimmunity in rheumatoid arthritis patients treated with infliximab	Design
Y.; Fujii Ishikawa, T.; Kondo-Ishikawa, S.; Hashimoto, M.; Furu, M.; Ito, H.; Imura, Y.; Yukawa, N.; Yoshifuji, H.; Ohmura, K.; Mimori, T.	Annals of the Rheumatic Diseases	Type i interferon plays a key role in immunogenicity and lupus-like autoimmunity in patients with rheumatoid arthritis treated by infliximab	Design
Y.; Fujii Ishikawa, T.; Ishikawa, S. K.; Yukawa, N.; Hashimoto, M.; Furu, M.; Ito, H.; Ohmura, K.; Mimori, T.	PLoS ONE [Electronic Resource]	Immunogenicity and Lupus-Like Autoantibody Production Can Be Linked to Each Other along With Type I Interferon Production in Patients with Rheumatoid Arthritis Treated With Infliximab: A Retrospective Study of a Single Center Cohort	Design

Authors	Source	Title	Reasons for exclusion
P.; Vinograi Isomaki, V.; Peltomaki, J.; Sokka-Isler, T.; Mali, M.; Vidqvist, K. L.; Haapala, A. M.; Korpela, M.; Makinen, H.	Annals of the Rheumatic Diseases	Therapeutic drug monitoring in arthritis patients receiving infliximab in daily clinical practice	Design
A.; Nurmohamed Jamnitski, M. T.; Hart, M. M.; Dijkmans, B. A.; Aarden, L.; Wolbink, G. J.	Arthritis and Rheumatism	Patients not responding to etanercept obtain lower trough etanercept concentrations compared to responding patients	Design
M.; Barton Jani, A.; Warren, R. B.; Griffiths, C. E. M.; Chinoy, H.	Rheumatology (United Kingdom)	The role of DMARDs in reducing the immunogenicity of TNF inhibitors in chronic inflammatory diseases	Design
M.; Chinoy Jani, H.; Warren, R. B.; Fu, B.; Griffiths, C. E.; Morgan, A. W.; Wilson, G.; Hyrich, K. L.; Isaacs, J. D.; Barton, A.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Influence of immunogenicity and drug levels on the efficacy of long-term treatment of rheumatoid arthritis with adalimumab and etanercept: A uk-based prospective study	Design
M.; Chinoy Jani, H.; Warren, R. B.; Griffiths, C. E. M.; Morgan, A. W.; Wilson, A. G.; Hyrich, K. L.; Isaacs, J.; Plant, D.; Barton, A.	Arthritis and Rheumatology	Clinical utility of random anti-TNF drug level testing and measurement of anti-drug antibodies on long-term treatment response in rheumatoid arthritis	Design
M.; Chinoy Jani, H.; Isaacs, J.; Morgan, A. W.; Wilson, A.; Hyrich, K. L.; Plant, D.; Barton, A.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Clinical utility and factors associated with certolizumab pegol drug levels and anti-drug antibodies in the long-term treatment of rheumatoid arthritis	Population
M.; Chinoy Jani, H.; Warren, R. B.; Griffiths, C. E.; Plant, D.; Fu, B.; Morgan, A. W.; Wilson, A. G.; Isaacs, J. D.; Hyrich, K.; Barton, A.; Biologics in Rheumatoid Arthritis, Genetics; Genomics Study Syndicate, Collaborators	Arthritis & Rheumatology	Clinical utility of random anti-tumor necrosis factor drug-level testing and measurement of antidrug antibodies on the long-term treatment response in rheumatoid arthritis.[Erratum appears in Arthritis Rheumatol. 2015 Nov;67(11):3096; PMID: 26508467]	Design

Authors	Source	Title	Reasons for exclusion
M.; Chinoy Jani, H.; Warren, R. B.; Griffiths, C. E.; Plant, D.; Morgan, A. W.; Wilson, A. G.; Hyrich, K. L.; Isaacs, J.; Barton, A.	Lancet	Clinical utility of random anti-tumour necrosis factor drug testing and measurement of anti-drug antibodies on long-term treatment response in rheumatoid arthritis	Design
M.; Chinoy Jani, H.; Warren, R. B.; Griffiths, C. E. M.; Plant, D.; Fu, B.; Morgan, A. W.; Wilson, A. G.; Isaacs, J. D.; Hyrich, K. L.; Barton, A.	Rheumatology	CLINICAL UTILITY OF RANDOM ANTI-TNF DRUG LEVEL TESTING AND MEASUREMENT OF ANTI-DRUG ANTIBODIES ON LONG-TERM TREATMENT RESPONSE IN RHEUMATOID ARTHRITIS	Design
M.; Isaacs Jani, J.; Morgan, A. W.; Wilson, A. G.; Plant, D.; Hyrich, K.; Chinoy, H.; Barton, A.	Rheumatology	HIGH FREQUENCY OF ANTI-DRUG ANTIBODIES AND CORRELATION OF LOW RANDOM DRUG LEVELS WITH LACK OF EFFICACY IN CERTOLIZUMAB PEGOL-TREATED PATIENTS WITH RHEUMATOID ARTHRITIS	Design
M.; Isaacs Jani, J. D.; Morgan, A. W.; Wilson, A. G.; Plant, D.; Hyrich, K. L.; Chinoy, H.; Barton, A.	Rheumatology	Detection of anti-drug antibodies using a bridging ELISA compared with radioimmunoassay in adalimumab-treated rheumatoid arthritis patients with random drug levels	Design
M.; Isaacs Jani, J. D.; Morgan, A. W.; Wilson, A. G.; Plant, D.; Hyrich, K. L.; Chinoy, H.; Barton, A.; Braggss,	Annals of the Rheumatic Diseases	High frequency of antidrug antibodies and association of random drug levels with efficacy in certolizumab pegol-treated patients with rheumatoid arthritis: results from the BRAGGSS cohort	Design
M.; Isaacs Jani, J. D.; Morgan, A. W.; Wilson, A. G.; Plant, D.; Hyrich, K. L.; Chinoy, H.; Barton, A.; Braggss,	Annals of the Rheumatic Diseases	High frequency of antidrug antibodies and OPEN ACCESS association of random drug levels with efficacy in certolizumab pegol-treated patients with rheumatoid arthritis: results from the BRAGGSS cohort	Design

Authors	Source	Title	Reasons for exclusion
E.; Garcia Jimenez, M.; De Guadiana, L. G.; Conesa, P.; Hernando, A.; De Bejar, A.; Pedregosa, J.; Vilchez, J. A.; Garcia, I.; Albaladejo, M. D.	Clinical Chemistry and Laboratory Medicine	Comparison of two different immunoassays to measure levels of infliximab and autoantibodies	Design
A.; Martinez-Feito Jochems, A.; Plasencia, C.; Hernandez-Breijo, B.; Mezcua, A.; Villalba, A.; Monjo, I.; Nozal, P.; Balsa, A.; Pascual-Salcedo, M. D.	Annals of the Rheumatic Diseases	Optimal circulating adalimumab levels range associated with good clinical response in rheumatoid arthritis patients	Design
K. K.; Goll Jorgensen, G. L.; Sexton, J.; Olsen, I. C.; Bolstad, N.; Lundin, K. E.; Berset, I. P.; Haavardsholm, E. A.; Mork, C.; Kvien, T. K.; Jahnsen, J.	Journal of Crohn's and Colitis	Long-term efficacy and safety of biosimilar infliximab (CT-P13) after switching from originator infliximab: Explorative subgroup analyses in IBD from the NOR-SWITCH EXTENSION trial	Population
P. D.; Juan Antonio Jose, V. A.; Irene, G. G.; Pablo, P. C.; Carlos, R. R.; Africa, D. B. A.; Ana, H. H.; Enrique Martin, J. S.; Iris, M. G.; Henar, G. L.; Ruben, M. T.; Maria Dolores, A. O.	Clinical Chemistry and Laboratory Medicine	Comparison of determination of adalimumab levels between two enzyme immunoassays (promonitor and sanquin)	Design
P. D.; Juan Antonio Jose, V. A.; Pablo, P. C.; Irene, G. G.; Carlos, R. R.; Ana, H. H.; Henar, G. L.; Enrique Martin, J. S.; Iris, M. G.; Africa, D. B. A.; Ruben, M. T.; Maria Dolores, A. O.	Clinical Chemistry and Laboratory Medicine	Comparison of determination of infliximab levels between two enzyme immunoassays (promonitor and sanquin)	Design
S. M.; Lee Jung, J. H.; Lee, J.; Suh, Y. S.; Koh, J. H.; Min, H. K.; Lee, J. Y.; Kwok, S. K.; Park, K. S.; Park, S. H.; Ju, J. H.	International Journal of Rheumatic Diseases	Immunogenicity of anti-TNF therapy in Korean patients with RA and AS	Design
T.; Plasencia Jurado, C.; Martin, S.; Navarro, R.; Bonilla, G.; Villalba, A.;	Annals of the Rheumatic Diseases	Comparison of golimumab levels detected by two different enzyme-linked immunosorbent assays: Promonitor vs sanquin	Design

Authors	Source	Title	Reasons for exclusion
Ramiro, S.; Jochems, A.; Balsa, A.; Pascual-Salcedo, D.			
T.; Plasencia Jurado, C.; Martinez-Feito, A.; Navarro-Compan, V.; Olariaga, E.; Diego, C.; Martin-Mola, E.; Balsa, A.; Pascual-Salcedo, D.	Annals of the Rheumatic Diseases	Low levels of infliximab at early stages predict the loss of drug levels and the clinical response at one year of treatment in patients with rheumatoid arthritis	Design
T.; Plasencia-Rodriguez Jurado, C.; Martinez, A.; Navarro-Compan, V.; Olariaga-Merida, E.; Peiteado, D.; Villalba, A.; Bonilla, G.; Diego, C.; Balsa, A.; Pascual-Salcedo, D.	Arthritis and Rheumatology	Infliximab low levels at early stages predict the loss of drug levels and the clinical response at one year of treatment in patients with rheumatoid arthritis	Design
G.; Czibula Kadar, A.; Szalay, B.; Nagy, K.; Pusztai, A.; Balog, A.; Monostori, E.; Vasarhelyi, B.; Szekanecz, Z.; Kovacs, L.	Annals of the Rheumatic Diseases	Predictors of disease course after the discontinuation of biologic therapy in rheumatoid arthritis patients with long-term remission	Design
J. R.; Schulze-Koops Kalden, H.	Nature Reviews Rheumatology	Immunogenicity and loss of response to TNF inhibitors: implications for rheumatoid arthritis treatment	Design
H. Kameda	Nippon Rinsho - Japanese Journal of Clinical Medicine	[TNF inhibitors]	Design
H. Kameda	Clinical Calcium	[Diagnosis and treatment of rheumatoid arthritis:toward the best practice. The best practice for TNF inhibitors.]	Design
J.; Chopra Kay, A.; Chandrashekara, S.; Olakkengil, D. J.; Bhojani, K. S.; Bhatia, G.; Rathi, G.; Thomas, M.; Maroli, S.; Thomson, E. S.; Shneyer, L.; Wyand, M. S.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	A phase 3, randomized, double-blind, active comparator study of the efficacy and safety of BOW015, a biosimilar infliximab, in patients with active rheumatoid arthritis on stable methotrexate doses	Design

Authors	Source	Title	Reasons for exclusion
P.; Hock Keating, B.; Barclay, M.; Stamp, L.; Spellerberg, M.; O'Donnell, J.	European Journal of Immunology	Application of an ELISA based competitive binding assay to measure concentration of anti-TNF biologics and neutralising anti-drug antibodies in the clinical laboratory	Design
M.; Codreanu Keiserman, C.; Handa, R.; Xibille-Friedmann, D.; Mysler, E.; Briceno, F.; Akar, S.	Expert Review of Clinical Immunology	The effect of antidrug antibodies on the sustainable efficacy of biologic therapies in rheumatoid arthritis: practical consequences	Design
J.; Drynda Kekow, S.	Arthritis and Rheumatology	Long persistence of anti-drug antibodies in adalimumab treated RA patients	Design
P. D. Kiely	Rheumatology	Biologic efficacy optimization--a step towards personalized medicine	Design
J. S.; Kim Kim, S. H.; Kwon, B.; Hong, S.	Expert Review of Clinical Immunology	Comparison of immunogenicity test methods used in clinical studies of infliximab and its biosimilar (CT-P13)	Design
E. L.; Pascual-Salcedo Kneepkens, D.; Plasencia, C.; Krieckaert, C. L. M.; Van Der Kleij, D.; Nurmohamed, M. T.; Lopez-Casla, M. T.; Rispens, T.; Wolbink, G.	Arthritis and Rheumatism	Golimumab levels, anti-drug antibodies and clinical response in rheumatoid arthritis patients at 28 week of follow-up	Design
E. L.; Plasencia Kneepkens, C.; Krieckaert, C. L.; Pascual-Salcedo, D.; van der Kleij, D.; Nurmohamed, M. T.; Lopez-Casla, M. T.; Wieringa, R.; Rispens, T.; Wolbink, G.	Annals of the Rheumatic Diseases	Golimumab trough levels, antidrug antibodies and clinical response in patients with rheumatoid arthritis treated in daily clinical practice	Population
E. L.; Pouw Kneepkens, M. F.; Wolbink, G. J.; Schaap, T.; Nurmohamed, M. T.; de Vries, A.; Rispens, T.; Bloem, K.	British Journal of Clinical Pharmacology	Dried blood spots from finger prick facilitate therapeutic drug monitoring of adalimumab and anti-adalimumab in patients with inflammatory diseases	Design

Authors	Source	Title	Reasons for exclusion
E. L.; Van Den Oever Kneepkens, I. A.; Plasencia, C.; Salcedo Pascual, D.; Lopez-Casla, M. T.; Van Der Kleij, D.; Nurmohamed, M. T.; Rispens, T.; Balsa, A.; Wolbink, G. J.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Tocilizumab levels are associated with clinical response in patients with rheumatoid arthritis	Intervention
E. L.; Wei Kneepkens, J. C. C.; Nurmohamed, M. T.; Yeo, K. J.; Chen, C. Y.; van der Horst-Bruinsma, I. E.; van der Kleij, D.; Rispens, T.; Wolbink, G.; Krieckaert, C. L. M.	Annals of the Rheumatic Diseases	Immunogenicity, adalimumab levels and clinical response in ankylosing spondylitis patients during 24 weeks of follow-up	Population
E.; van den Oever Kneepkens, I. A. M.; Plasencia, C. H.; Pascual-Salcedo, D.; de Vries, A.; Hart, M.; Nurmohamed, M. T.; Balsa, A.; Rispens, T.; Wolbink, G.	Scandinavian Journal of Rheumatology	Serum tocilizumab trough concentration can be used to monitor systemic IL-6 receptor blockade in patients with rheumatoid arthritis: a prospective observational cohort study	Intervention
Y.; Otal Koyama, T.; Miura, T.	Annals of the Rheumatic Diseases	Analysis of patients with detectable trough serum levels of infliximab revealed significant predictors associated with non-response to "actual infliximab in rheumatoid arthritis	Design
A. Kozmar	Biochemia Medica	The role of laboratories in optimizing biological therapy	Population
C.; Rispens Krieckaert, T.; Wolbink, G.	Current Opinion in Rheumatology	Immunogenicity of biological therapeutics: From assay to patient	Design
C.; Vogelzang Krieckaert, E.; Pouw, M.; Nurmohamed, M.; Wolbink, G.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Adalimumab serum concentrations in patients with rheumatoid arthritis or psoriatic arthritis taking concomitant DMARD therapy	Population
B.; King Kuang, L.; Wang, H. F.	Bioanalysis	Therapeutic monoclonal antibody concentration monitoring: Free or total?	Design

Authors	Source	Title	Reasons for exclusion
J.; Jokiranta Laine, T. S.; Eklund, K. K.; Vakevainen, M.; Puolakka, K.	Biologics	Cost-effectiveness of routine measuring of serum drug concentrations and anti-drug antibodies in treatment of rheumatoid arthritis patients with TNF-alpha blockers	Design
D.; Wong Langguth, P.; Bowling, A.; Bagga, H.; Freeman, D.; Ford, E.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Serum trough levels of adalimumab inversely correlate with disease activity in patients with inflammatory arthritis	Population
Y.; Youssef Leow, P.; Richards, B.	International Journal of Rheumatic Diseases	Correlation of adalimumab trough level with disease activity in patients with inflammatory arthritides	Population
J. H.; Xu Leu, Z.; Hu, C.; Mendelsohn, A.; Ford, J.; Davis, H. M.; Zhou, H.	Arthritis and Rheumatism	Importance of Steady-State Trough Concentrations After Intravenous Golimumab with Concomitant Methotrexate in Subjects with Active Rheumatoid Arthritis	Design
M. H.; Li Li, H. Z.; Gao, K.; Wang, M. Y.; An, W. Q.; Zhu, Y. R.; Ding, L.; Wang, L.; Gu, J. L.; Zuo, G. L.; Sun, L.	Journal of Immunological Methods	A simple and cost-effective assay for measuring anti-drug antibody in human patients treated with Adalimumab	Design
F.; de Salazar Llinares-Tello, J. R. G.; Senabre-Gallego, J. M.; Santos-Soler, G.; Santos-Ramirez, C.; Salas-Heredia, E.; Barber-Valles, X.; Molina-Garcia, J.; Aire-Mb Grp	Rheumatology International	Practical application of acid dissociation in monitoring patients treated with adalimumab	Design
F.; de Salazar Llinares-Tello, J. R. G.; Senabre-Gallego, J. M.; Santos-Soler, G.; Santos-Ramirez, C.; Salas-Heredia, E.; Molina-Garcia, J.; Aire-Mb Grp	Clinical Chemistry and Laboratory Medicine	Analytical and clinical evaluation of a new immunoassay for therapeutic drug monitoring of etanercept	Design
F.; Rosas Llinares-Tello, J.; De La Torre, I.; Valor, L.; Senabre, J. M.; Barber, X.; Hernandez, D.; Carreno, L.;	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Comparative study of both versions of an immunoassay commercialized for therapeutic drug monitoring of adalimumab	Design

Authors	Source	Title	Reasons for exclusion
Santos-Soler, G.; Salas, E.; Santos-Ramirez, C.; Sanchez-Barrioluengo, M.; Molina-Garcia, J.			
F.; Rosas Llinares-Tello, J.; de la Torre, I.; Valor, L.; Barber, X.; Senabre, J. M.; el Grupo Aire-Mb, Hugm	Reumatologia Clinica	Comparative study of both versions of an immunoassay commercialized for therapeutic drug monitoring of adalimumab in rheumatoid arthritis	Design
F.; Rosas Llinares-Tello, J.; Senabre-Gallego, J. M.; Molina, J.; Salas, E.; Santos-Soler, G.; Santos Ramirez, C.; Ortega, R.; Barber, X.; Pons, A.; Cano, C.; Lorente, M.; Sanchez-Barrioluengo, M.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Usefulness of the acid dissociation in inmunogenicity detection in patients in treatment with anti-TNF drugs	Design
F.; Rosas Llinares-Tello, J.; Senabre-Gallego, J. M.; Santos-Soler, G.; Santos-Ramirez, C.; Salas-Heredia, E.; Barber, X.; Molina, J.; Cano, C.; Pons, A.	Arthritis and Rheumatology	Implementation of an acid dissociation procedure for immunogenicity detection in patients treated with anti-TNF drugs	Design
F.; Rosas-Gomez de Salazar Llinares-Tello, J.; Senabre-Gallego, J. M.; Santos-Soler, G.; Santos-Ramirez, C.; Salas-Heredia, E.; Barber-Valles, X.; Molina-Garcia, J.; Aire-Mb Group	Rheumatology International	Practical application of acid dissociation in monitoring patients treated with adalimumab.[Erratum appears in Rheumatol Int. 2014 Dec;34(12):1709]	Design
J. A.; Golikova Lopatnikova, E. A.; Shkaruba, N. S.; Sizikov, A. E.; Sennikov, S. V.	Scandinavian Journal of Rheumatology	Analysis of the levels of tumour necrosis factor (TNF), autoantibodies to TNF, and soluble TNF receptors in patients with rheumatoid arthritis	Design
R.; Martinez Lopez-Rodriguez, A.; Plasencia, C.; Jochems, A.; Pascual-Salcedo, D.; Balsa, A.; Gonzalez, A.	Annals of the Rheumatic Diseases	Increased frequency of anti-drug antibodies in patients carrying compatible IgG1 allotypes and treated with anti-TNF antibodies	Design

Authors	Source	Title	Reasons for exclusion
G.; Sigidin Lukina, Y.; Alexandrova, E.; Novikov, A.; Aronova, E.; Kanonirova, M.; Glukhova, S.; Nasonov, E.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Clinical significance of antibodies to infliximad in rheumatoid arthritis (RA) patients	Population
E.; Vultaggio Maggi, A.; Matucci, A.	Expert Review of Clinical Immunology	Acute infusion reactions induced by monoclonal antibody therapy	Design
P.; Real Maid, R.; Pedersen, R.; Shen, Q.; Hidalgo, R.	Journal of Clinical Rheumatology	Incidence of anti-drug antibodies in patients with rheumatoid arthritis from argentina treated with adalimumab, etanercept, or infliximab in a real-world setting	Design
J. R.; Salgado Maneiro, E.; Gomez-Reino, J. J.	JAMA Internal Medicine	Immunogenicity of monoclonal antibodies against tumor necrosis factor used in chronic immune-mediated Inflammatory conditions: systematic review and meta-analysis	Design
B.; Botti Marinari, E.; Bavetta, M.; Spallone, G.; Zangrilli, A.; Talamonti, M.; Richetta, A.; Chimenti, S.; Costanzo, A.	Drug Development Research	Detection of adalimumab and anti-adalimumab levels by ELISA: clinical considerations	Population
H.; Maslinski Marotte, W.; Miossec, P.	Arthritis Research & Therapy	Circulating tumour necrosis factor-alpha bioactivity in rheumatoid arthritis patients treated with infliximab: link to clinical response	Population
H.; Rinaudo Marotte, M.; Paul, S.; Fautrel, B.	Annals of the Rheumatic Diseases	No prediction of relapse by TNF blocker concentrations or detection of antibodies against anti-TNF: Data from strass study	Design
H.; Rinaudo-Gaujous Marotte, M.; Paul, S.; Fautrel, B.	Arthritis and Rheumatology	TNF blocker concentrations or detection of antibodies against anti-TNF before a tapering process are not predictive to relapse	Design

Authors	Source	Title	Reasons for exclusion
A.; L'Ami Marsman, M.; Kneepkens, E.; Kienhorst, L.; Nurmohamed, M.; Krieckaert, C.; Wolbink, G.	Annals of the Rheumatic Diseases	Patient reported reasons for refraining from participation in dose reduction studies with biologics	Design
L.; Olivera Martelli, P.; Roblin, X.; Attar, A.; Peyrin-Biroulet, L.	Journal of Gastroenterology	Cost-effectiveness of drug monitoring of anti-TNF therapy in inflammatory bowel disease and rheumatoid arthritis: a systematic review	Design
S.; Del Agua Martin, A. R.; Torres, N.; Pascual-Salcedo, D.; Plasencia, C.; Jurado, T.; Ruiz- Arguello, B.; Martinez, A.; Navarro, R.; Nagore, D.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Validation and comparison study of immunoassays for the measurement of golimumab and antibodies to golimumab in rheumatic patients	Design
L. P.; Valor Martinez Estupinan, L.; Hernandez, D.; Naredo, E.; Montoro, M.; Nieto-Gonzalez, J. C.; Mata-Martinez, C.; Ovallez-Bonilla, J.; Serrano-Benavente, B.; Gonzalez-Fernandez, C.; Lopez-Longo, J.; Monteagudo, I.; Carreno-Perez, L.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Relation between serum infliximab levels and changes of rheumatoid factor and antibodies to citrullinated peptides levels in patients with rheumatoid arthritis	Population
L.; Hernandez-Florez Martinez-Estupinan, D.; Janta, I.; Ovalles-Bonilla, J. G.; Nieto, J. C.; Gonzalez-Fernandez, C. M.; Del Rio, T.; Monteagudo, I.; Lopez-Longo, F. J.; Naredo, E.; Valor, L.	Clinical & Experimental Rheumatology	An exploratory study to determine whether infliximab modifies levels of rheumatoid factor and antibodies to cyclic citrullinated peptides in rheumatoid arthritis patients	Design
A.; Bravo Gallego Martinez-Feito, L. Y.; Hernandez-Breijo, B.; Plasencia, C.; Jochems, A.; Gonzalez, M. A.; Monjo, I.; Peiteado, D.; Bonilla, G.; Nozal, P.; Balsa, A.; Pascual-Salcedo, D.	Annals of the Rheumatic Diseases	Clinical relevance of detecting anti-adalimumab antibodies with a drug-tolerant assay	Design

Authors	Source	Title	Reasons for exclusion
A.; Plasencia Martinez-Feito, C.; Villalba, A.; Jurado, T.; Mezcua, A.; Martin-Mola, E.; Bonilla, G.; Balsa, A.; Pascual-Salcedo, D.	Annals of the Rheumatic Diseases	Effect of methotrexate in the presence of drug and the appearance of antibodies against TNF inhibitors in patients with rheumatoid arthritis	Design
M.; Carmona Martin-Lopez, L.; Balsa, A.; Calvo-Alen, J.; Sanmarti, R.; Tornero, J.; Rosas, J.	Rheumatology International	Serum drug levels of biologic agents in the management of rheumatoid arthritis and spondyloarthritis: a systematic review	Design
Y.; Narazaki Matsuura, M.; Nishide, M.; Kato, Y.; Yorifuji, H.; Hirano, T.; Shima, Y.; Tanaka, T.; Ogata, A.; Kumanogoh, A.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Optimization of treatment intervals of tocilizumab and golimumab by measuring serum trough levels in rheumatoid arthritis patients	Design
A.; Petroni Matucci, G.; Nencini, F.; Pratesi, S.; Maggi, E.; Vultaggio, A.	Allergy: European Journal of Allergy and Clinical Immunology	Anti-infliximab antibodies production and clinical consequences: Adverse reactions and loss of response	Population
A.; Vultaggio Matucci, A.; Nencini, F.; Pratesi, S.; Rossi, O.; Parronchi, P.; Romagnani, S.; Maggi, E.	Allergy: European Journal of Allergy and Clinical Immunology	Adverse reactions to biological agents: Role of anti-infliximab antibodies and analysis of potential risk factors	Design
D.; Gainaru Mazilu, C.; Apetrei, N.; Luca, G.; Gudu, T.; Peltea, A.; Constantinescu, C.; Saulescu, I.; Bojinca, V.; Balanescu, A.; Predeteanu, D.; Ionescu, R.; Opris, D.	International Journal of Rheumatic Diseases	Methotrexate and Infliximab immunogenicity	Design
D.; Opris Mazilu, D.; Gainaru, C.; Iliuta, M.; Apetrei, N.; Luca, G.; Borangiu, A.; Gudu, T.; Peltea, A.; Groseanu, L.; Constantinescu, C.; Saulescu, I.; Bojinca, V.; Balanescu, A.; Predeteanu, D.; Ionescu, R.	BioMed Research International	Monitoring drug and antidrug levels: a rational approach in rheumatoid arthritis patients treated with biologic agents who experience inadequate response while being on a stable biologic treatment	Population
D.; Opris Mazilu, D.; Iachim, E.; Deaconu, C.; Saulescu, I.; Borangiu,	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of	Time to first signs of loss of response in rheumatoid arthritis patients treated with time to	Design

Authors	Source	Title	Reasons for exclusion
A.; Grosanu, L.; Constantinescu, C.; Balanescu, A.; Predeteanu, D.; Ionescu, R.	Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	first signs of loss of response in rheumatoid arthritis patients treated with anti TNF agents: Correlations with serum drug level, immunogenicity and csDMARD association	
F.; Plasencia Medina, C.; Goupille, P.; Ternant, D.; Balsa, A.; Mulleman, D.	Therapeutic Drug Monitoring	Current Practice for Therapeutic Drug Monitoring of Biopharmaceuticals in Rheumatoid Arthritis	Design
J. C.; Mulleman Meric, D.; Paintaud, G.; Ducourau, E.; Magdelaine-Beuzelin, C.; Valat, J. P.; Goupille, P.	Arthritis and Rheumatism	Infliximab concentration monitoring improves the control of disease activity in rheumatoid arthritis	Design
P. L.; Valentini Meroni, G.; Ayala, F.; Cattaneo, A.; Valesini, G.	Autoimmunity Reviews	New strategies to address the pharmacodynamics and pharmacokinetics of tumor necrosis factor (TNF) inhibitors: A systematic analysis	Design
P.; Charlotte Mieke, K.; Michael, N.; Margreet, H.; Henk, T. V.; Desiree, V. D. K.; Lucien, A.; Theo, R.; Gertjan, W.	Clinical Chemistry and Laboratory Medicine	Measurement of anti-TNF drugs levels is the key to optimal, personalized and cost-effective treatment	Population
V. I.; Cavalier Mistretta, E.; Collette, J.; Lutteri, L.; Chapelle, J. P.	Revue Medicale de Liege	Interest of monoclonal antibodies in the biomedical laboratory analysis. [French]	Design
T.; Momohara Mochizuki, S.; Ikari, K.; Okamoto, H.; Kobayashi, S.; Tsukahara, S.; Iwamoto, T.; Kawamura, K.; Saito, S.; Tomatsu, T.	Modern Rheumatology	The serum concentration of infliximab in cases of autologous blood donation for patients with rheumatoid arthritis	Design
C. C.; Fong Mok, B.; Ho, L. Y.; To, C. H.	Annals of the Rheumatic Diseases	Serum levels of the anti-TNF biologics correlate with clinical efficacy in patients with inflammatory arthritis	Design

Authors	Source	Title	Reasons for exclusion
C. C.; Fong Mok, L. S.; Ho, L. Y.; To, C. H.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Serum levels of the anti-TNF biologics correlate with clinical efficacy in patients with inflammatory arthritis	Design
C. C.; Tsai Mok, W. C.; Chen, D. Y.; Wei, J. C.	Expert Opinion on Biological Therapy	Immunogenicity of anti-TNF biologic agents in the treatment of rheumatoid arthritis	Design
C. C.; Van Der Kleij Mok, D.; Wolbink, G.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Anti-drug antibodies, drug levels and clinical efficacy of the anti-TNF biologics in rheumatic diseases	Population
C. C.; van der Kleij Mok, D.; Wolbink, G. J.	Clinical Rheumatology	Drug levels, anti-drug antibodies, and clinical efficacy of the anti-TNFalpha biologics in rheumatic diseases	Population
R. J.; Xavier Moots, R.; Mok, C. C.; Rahman, M. U.; Tsai, W. C.; Al Maini, M.; Pavelka, K.; Mahgoub, E.; Kotak, S.; Korth-Bradley, J.; Pedersen, R.; Mele, L.; Shen, Q.; Vlahos, B.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Incidence of anti-drug antibodies in rheumatoid arthritis patients treated with adalimumab, etanercept, or infliximab in a real-world setting	Population
R. J.; Xavier Moots, R. M.; Mok, C. C.; Rahman, M. U.; Tsai, W. C.; Al-Maini, M. H.; Pavelka, K.; Mahgoub, E.; Kotak, S.; Korth-Bradley, J.; Pedersen, R.; Mele, L.; Shen, Q.; Vlahos, B.	PLoS ONE [Electronic Resource]	The impact of anti-drug antibodies on drug concentrations and clinical outcomes in rheumatoid arthritis patients treated with adalimumab, etanercept, or infliximab: Results from a multinational, real-world clinical practice, non-interventional study.[Erratum appears in PLoS One. 2017 Jun 5;12 (6):e0179308; PMID: 28582423]	Population
S. Mori	Modern Rheumatology	A relationship between pharmacokinetics (PK) and the efficacy of infliximab for patients with rheumatoid arthritis: characterization of infliximab-resistant cases and PK-based modified therapy	Design

Authors	Source	Title	Reasons for exclusion
S.; Ueki Mori, Y.	Modern Rheumatology	Primary lack of efficacy of infliximab therapy for rheumatoid arthritis: pharmacokinetic characterization and assessment of switching to tocilizumab	Design
D.; Ducourau Mulleman, E.; Paintaud, G.; Ternant, D.; Watier, H.; Goupille, P.	Joint Bone Spine	Should anti-TNF-alpha drug levels and/or anti-drug antibodies be assayed in patients treated for rheumatoid arthritis?	Design
D.; Lin Mulleman, D. C. M.; Ducourau, E.; Emond, P.; Ternant, D.; Magdelaine-Beuzelin, C.; Valat, J. P.; Paintaud, G.; Goupille, P.	Therapeutic Drug Monitoring	Trough Infliximab Concentrations Predict Efficacy and Sustained Control of Disease Activity in Rheumatoid Arthritis	Design
D.; Meric Mulleman, J. C.; Paintaud, G.; Ducourau, E.; Magdelaine-Beuzelin, C.; Valat, J. P.; Goupille, P.	Arthritis Research and Therapy	Infliximab concentration monitoring improves the control of disease activity in rheumatoid arthritis	Population
J.; Stamenkovic Nedovic, B.; Stojanovic, S.; Zivkovic, V.	Annals of the Rheumatic Diseases	Does concentration of antibodies to etanercept and adalimumab correlates with parameters of disease activity in patients with rheumatoid arthritis?	Population
K.; Hashizume Nishida, K.; Kadota, Y.; Natsumeda, M.; Nakahara, R.; Saito, T.; Kanazawa, T.; Ezawa, K.; Ozaki, T.	Modern Rheumatology	Time-concentration profile of serum etanercept in Japanese patients with rheumatoid arthritis after treatment discontinuation before orthopedic surgery	Design
A.; Garces Nunes, S.; Vieira, A.; Demangeot, J.; Freitas, J.	Journal of Crohn's and Colitis	Infliximab trough levels and anti-infliximab antibodies in rheumatoid arthritis and in IBD patients A comparison from a single center	Population
J.; Liu O'Donnell, J.; Keating, P.; Hock, B.; Spellerberg, M.; Barclay, M.; Stamp, L.	Internal Medicine Journal	Anti-drug antibodies (ADA): Assay performance in patients treated for inflammatory bowel and rheumatic disease with biodrugs, adalimumab and infliximab	Design

Authors	Source	Title	Reasons for exclusion
M.; Tercelj Ogric, M.; Praprotnik, S.; Tomsic, M.; Bozic, B.; Sodin-Semrl, S.; Cucnik, S.	Immunologic Research	Detection of adalimumab and anti-adalimumab antibodies in patients with rheumatoid arthritis: a comprehensive overview of methodology pitfalls and benefits	Design
F.; Beggio Ometto, M.; Friso, L.; Astorri, D.; Raffeiner, B.; Botsios, C.; Bernardi, L.; Padoan, R.; Punzi, L.; Ghiraldello, A.; Doria, A.	Annals of the Rheumatic Diseases	Anti-etanercept antibodies and etanercept leves levels in rheumatoid arthritis patients treated with low and full-dose etanercept in DAS28 remission	Design
D.; Borangiu Opris, A.; Gudu, T.; Mazilu, D.; Balanescu, A.; Saulescu, I.; Ionescu, R.	Annals of the Rheumatic Diseases	Does serum drug level correlates with ultrasound evaluation in patients with rheumatoid arthritis treated with TNF antagonists?	Design
D.; Diana Opris, M.; Gainaru, C.; Iliuta, M.; Groseanu, L.; Saulescu, I.; Constantinescu, C.; Bojinca, V.; Balanescu, A.; Predeteanu, D.; Ionescu, R.	Annals of the Rheumatic Diseases	SERUM DRUG LEVEL AND ANTI-CITRULLINATED PEPTIDE ANTIBODIES AS BIOMARKERS THAT PREDICT EULAR RESPONSE IN RHEUMATOID ARTHRITIS - A NEW STEP TO PERSONALIZED MEDICINE	Intervetion
D.; Mazilu Opris, D.; Bojinca, V.; Balanescu, A.; Borangiu, A.; Ionescu, R.	Clinical and Experimental Rheumatology	Adalimumab serum drug level correlates to clinical response in patients with rheumatoid arthritis	Population
D.; Mazilu Opris, D.; Bojinca, V.; Saulescu, I.; Balanescu, A.; Ionescu, R. M.	Clinical and Experimental Rheumatology	Secondary failure to etanercept in rheumatoid arthritis patients-the role of immunogenicity, characteristics andevolution of the disease	Population
D.; Mazilu Opris, D.; Ionescu, R.	Clinical and Experimental Rheumatology	Clinical response in rheumatoid arthritis patients with anti-infliximab antibodies	Population
A.; Padulles Padulles, N.; Lloberas-Blanch, N.; Juanola, X.; Narvaez, F. J.; Leiva, E.; Cobo, S.; Bas, J.; Climent, J.; Carrere, M.; Colom, H.	European Journal of Hospital Pharmacy	Evaluation of a population pharmacokinetic model of infliximab in rheumatoid arthritis for prediction of individual dosage requirements	Design

Authors	Source	Title	Reasons for exclusion
R.; Schmitt Palaparthy, S.; Rehman, M. I.; Cai, C. H.; Wang, K.; Von Richter, O.	Journal of Crohn's and Colitis	Incidence and impact of immunogenicity in a randomised, double-blind phase III study comparing a proposed infliximab biosimilar (PF-06438179/GP1111) with reference infliximab	Population
M. A.; Purushothama Partridge, S.; Elango, C.; Lu, Y. M.	Journal of Immunology Research	Emerging Technologies and Generic Assays for the Detection of Anti-Drug Antibodies	Design
V.; De Santis Pecoraro, E.; Melegari, A.; Trenti, T.	Autoimmunity Reviews	The impact of immunogenicity of TNF alpha inhibitors in autoimmune inflammatory disease. A systematic review and meta-analysis	Design
A. Perdriger	Biologics	Infliximab in the treatment of rheumatoid arthritis	Design
G.; Pratesi Petroni, S.; Nencini, F.; Milla, M.; Maggi, E.; Matucci, A.; Vultaggio, A.	Allergy: European Journal of Allergy and Clinical Immunology	The onset of anti-infliximab antibodies occurs after the first drug infusions and their high levels are related to adverse reactions	Population
A.; Pascual-Salcedo Pieren, D.; Aguado, P.; Bonilla, G.; De Miguel, E.; Monjo, I.; Nuno, L.; Peiteado, D.; Villalba, A.; Coro, E. M.; Tornero, C.; Bogas, P.; Balsa, A.; Plasencia-Rodriguez, C.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Flare incidence and predictive factors in a population of patients with rheumatoid arthritis under optimised treatment with adalimumab and infliximab	Population
C.; Jurado Plasencia, T.; Villalba, A.; Peitedado, D.; Casla, M. T.; Nuno, L.; Bonilla, M. G.; Martinez-Feito, A.; Martin-Mola, E.; Pascual-Salcedo, D.; Balsa, A.	Frontiers in Medicine	Effect of Infliximab Dose Increase in Rheumatoid Arthritis at Different Trough Concentrations: A Cohort Study in Clinical Practice Conditions	Population
C.; Pascual-Salcedo Plasencia, D.; Alcozer, P.; Garcia-Carazo, S.; Franco, K. N.; Cajigas, D.; Bonilla, G.; Lojo, L.; Nuno, L.; Villalba, A.; Peiteado, D.;	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Etanercept serum trough levels are correlated with clinical activity in rheumatoid arthritis patients with long-term treatment with etanercept	Population

Authors	Source	Title	Reasons for exclusion
Arribas, F.; Lopez-Casla, M. T.; Martin-Mola, E.; Balsa, A.			
C.; Pascual-Salcedo Plasencia-Rodriguez, D.; Bonilla, M. G.; Villalba, A.; Peiteado, D.; Nuno, L.; Aguado, P.; Jurado, T.; Martin-Mola, E.; Balsa, A.	Arthritis and Rheumatology	The monitoring of infliximab levels at early stages can predict the development of anti-infliximab antibodies in a cohort of rheumatoid arthritis patients treated with infliximab	Outcome
C.; Pascual-Salcedo Plasencia-Rodriguez, M. D.; Bonilla, G.; Navarro-Compan, V.; Martinez-Feito, A.; Diego, C.; Villalba, A.; Peiteado, D.; Nuno, L.; Martin-Mola, E.; Balsa, A.	Annals of the Rheumatic Diseases	Influence of drug levels during the first anti-TNF therapy on the clinical response to a second biologic in rheumatoid arthritis patients	Intervention
M. F.; Krieckaert Pouw, C. L.; Nurmohamed, M. T.; Rispens, T.; Aarden, L.; Wolbink, G.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Adalimumab trough level in blood corresponding with clinical response	Population
M. F.; Krieckaert Pouw, C. L.; Nurmohamed, M. T.; van der Kleij, D.; Aarden, L.; Rispens, T.; Wolbink, G.	Annals of the Rheumatic Diseases.	Key findings towards optimising adalimumab treatment: The concentration-effect curve	Intervention
M. F.; Mulleman Pouw, D.; Nurmohamed, M. T.; Rispens, T.; Paintaud, G.; Wolbink, G.; Ternant, D.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Adalimumab concentration at 16 weeks of treatment is associated with treatment discontinuation within one year	Intervention
M. S.; Bendtzen Prado, K.; Andrade, L. E. C.	Expert Opinion on Drug Metabolism and Toxicology	Biological anti-TNF drugs: immunogenicity underlying treatment failure and adverse events	Design
L. Puig	Journal of the European Academy of Dermatology and Venereology	Defining effective approaches to the reduction or elimination of biologic therapy immunogenicity and loss of response	Population

Authors	Source	Title	Reasons for exclusion
E.; Alvarez-De La Sierra Quesada-Masachs, D.; Garcia Prat, M.; Pujol-Borrell, R.; Martinez Gallo, M.; Modesto Caballero, C.; Marin Sanchez, A. M.	Annals of the Rheumatic Diseases	Prospective analysis of the immunogenic response in JIA patients (paediatric and adult) on antiTNF treatment	Population
Trdj; Svenson Radstake, M.; Eijsbouts, A. M.; van den Hoogen, F. H. J.; Enevold, C.; van Riel, Plcm; Bendtzen, K.	Annals of the Rheumatic Diseases	Formation of antibodies against infliximab and adalimumab strongly correlates with functional drug levels and clinical responses in rheumatoid arthritis	Population
B.; Delgado Reyes-Beltran, G.	Journal of Immunotoxicology	Anti-drug antibodies in Colombian patients with rheumatoid arthritis treated with Enbrel vs Etanar - Preliminary report	Population
S.; Martinez-Morillo Rodriguez-Muguruza, M.; Sanint, J.; Quirant Sr, B.; Teniente Sr, A.; Prior, A.; Riveros-Frutos, A.; Holgado, S.; Mateo, M. L.; Olive, A.; Canellas, J.; Tena, X.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Tocilizumab serum levels and antidrug antibodies and its relationship with disease activity in rheumatic diseases	Intervention
P. N.; Mignot Roland, S. G.; Bruns, A.; Hurtado, M.; Palazzo, E.; Hayem, G.; Dieude, P.; Meyer, O.; Martin, S. C.	Arthritis Research and Therapy	Antibodies to mutated citrullinated vimentin for diagnosing rheumatoid arthritis in anti-CCP-negative patients and for monitoring infliximab therapy	Intervention
J.; L. Linares F; De La Torre Rosas, I.; Valor, L.; Barber, X.; Santos-Ramirez, C.; Hernandez, D.; Senabre, J. M.; Carreno, L.; Santos-Soler, G.; Salas, E.; Sanchez, BArrioluengo M.; Molina-Garcia, J.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Clinical usefulness of serum level of adalimumab, in patients with rheumatoid arthritis	Outcome
J.; Llinares-Tello Rosas, F.; Senabre, J. M.; Santos-Ramirez, C.; Santos-Soler, G.; Salas, E.; Barber, X.; Sanchez-	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Evaluation of anti-TNF levels and anti-TNF antibodies in rheumatic diseases treated with infliximab and adalimumab; results from a local registry	Population

Authors	Source	Title	Reasons for exclusion
Barrioluengo, M.; Molina-Garcia, J.; Llahi, N.; Cano, C.			
J.; Llinares-Tello Rosas, F.; de la Torre, I.; Santos-Ramirez, C.; Senabre-Gallego, J. M.; Valor, L.; Barber-Valles, X.; Hernandez-Florez, D.; Santos-Soler, G.; Salas-Heredia, E.; Carreno, L.; Aire-Mb Grp	Clinical and Experimental Rheumatology	Clinical relevance of monitoring serum levels of adalimumab in patients with rheumatoid arthritis in daily practice	Design
J.; Llinares-Tello Rosas, F.; Martin, S.; Senabre, J. M.; Salas, E.; Oliver, S.; Santos Soler, G.; Santos Ramirez, C.; Barber, X.; Pons, A.; Cano, C.; Lorente, M.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Evaluation of serum level of golimumab and antibodies anti-golimumab in patients with rheumatic diseases: Results from a local registry	Population
B.; Maguregui Ruiz-Arguello, A.; Del Agua, A. R.; Pascual-Salcedo, D.; Jurado, T.; Plasencia, C.; Balsa, A.; Llinares-Tello, F.; Rosas, J.; Torres, N.; Martinez, A.; Nagore, D.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Antibodies to infliximab in remicade-treated rheumatic patients show identical reactivity towards biosimilars	Outcome
M. B.; Maguregui Ruiz-Arguello, A.; del Agua, A. R.; Pascual-Salcedo, D.; Martinez-Feito, A.; Jurado, T.; Plasencia, C.; Balsa, A.; Llinares-Tello, F.; Rosas, J.; Torres, N.; Martinez, A.; Nagore, D.	Annals of the Rheumatic Diseases	Antibodies to infliximab in Remicade-treated rheumatic patients show identical reactivity towards biosimilars	Outcome
V.; Bastida Ruiz-Esquide, C.; Pascal, M.; Yague, J.; Soy, D.; Sanmarti, R.	Annals of the Rheumatic Diseases	Therapeutic drug monitoring on rheumatoid arthritis patients with reduced doses of intravenous tocilizumab	Intervention
V.; Gonzalez-Navarro Ruiz-Esquide, A.; Yague, J.; Inciarte- Mundo, J.; Hernandez, M. V.; Ramirez, J.;	Arthritis and Rheumatology	Tocilizumab serum trough levels and its relationship with disease activity and drug dosage in rheumatoid arthritis patients	Intervention

Authors	Source	Title	Reasons for exclusion
Cabrera-Villalba, S.; Canete, J. D.; Sanmarti, R.			
V.; Gonzalez-Navarro Ruiz-Esquide, A.; Yague, J.; Ramirez, J.; Hernandez, M. V.; Cabrera-Villalba, S.; Inciarte-Mundo, J.; Canete, J. D. D.; Sanmarti, R.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Serum levels of tocilizumab and its relationship with disease activity and drug dosage in patients with rheumatoid arthritis	Intervention
V.; Zufferey Ruiz-Esquide, P.; Inciarte-Mundo, J.; Yague, J.; Hernandez, M. V.; Ramirez, J.; Berner, J.; Pascal, M.; Cuervo, A.; Canete, J. D.; Sanmarti, R.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Tocilizumab serum trough levels and disease activity in rheumatoid arthritis	Intervention
V.; Zufferey Ruiz-Esquide, P.; Yague, J.; Berner, J.; Inciarte-Mundo, J.; Gonzalez-Navarro, A.; Hernandez, V.; Ramirez, J.; Cuervo, A.; Canete, J.; Sanmarti, R.	Annals of the Rheumatic Diseases	Relationship between clinical remission and serum levels of tocilizumab in the treatment of rheumatoid arthritis	Intervention
P.; Vermeire Rutgeerts, S.; Van Assche, G.	Gut	Predicting the response to infliximab from trough serum levels	Population
T.; Nishida Saito, K.; Hashizume, K.; Nakahara, R.; Kanazawa, T.; Kadota, Y.; Ozaki, T.	International Journal of Rheumatic Diseases	Time-concentration profile of etanercept in the serum from Japanese rheumatoid arthritis patients after discontinuation before orthopaedic surgery	Population
R.; Inciarte Sanmarti, J.; Estrada Alarcon, P.; Garcia Manrique, M.; Gonzalez Navarro, A.; Narvaez, J.; Rodriguez-Moreno, J.; Gomez-Centeno, A.; Yague, J.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Immunogenicity of anti-TNF antagonists in patients with rheumatoid arthritis or polyarticular psoriatic arthritis in clinical remission or low disease activity: The inmunoremar study	Outcome
R.; Inciarte-Mundo Sanmarti, J.; Estrada Alarcon, P.; Garcia Manrique,	Annals of the Rheumatic Diseases	Serum levels of TNF antagonists in rheumatoid arthritis: Can we establish an optimal cut-off to	Design

Authors	Source	Title	Reasons for exclusion
M.; Narvaez, J.; Rodriguez, J.; Gomez Centeno, T.; Pascal, M.; Yague, J.		identify patients in remission or low disease activity?	
R.; Inciarte-Mundo Sanmarti, J.; Estrada-Alarcon, P.; Garcia-Manrique, M.; Narvaez, J.; Rodriguez-Moreno, J.; Gomez-Centeno, A.; Pascal, M.; Yague, J.	Annals of the Rheumatic Diseases	Towards optimal cut-off trough levels of adalimumab and etanercept for a good therapeutic response in rheumatoid arthritis. Results of the INMUNOREMAR study	Outcomes
R.; Inciarte-Mundo Sanmarti, J.; Estrada-Alarcon, P.; Garcia-Manrique, M.; Narvaez, J.; Gomez-Centeno, A.; Rodriguez-Moreno, J.; Pascal, M.; Yague, J.	Annals of the Rheumatic Diseases	Immunogenicity of TNF inhibitors in patients with rheumatoid arthritis or polyarticular psoriatic arthritis in clinical remission or low disease activity: A one-year multicentre prospective study (the inmunoremar study)	Design
M.; Takemura Sato, M.; Tani, T.; Ohashi, T.	Annals of the Rheumatic Diseases	Can infliximab efficacy be predicted based on blood concentration at the fourth dose?	Design
E. M. H.; Benoy-De Keuster Schmitz, S.; Meier, A. J. L.; Scharnhorst, V.; Traksel, R. A. M.; Broeren, M. A. C.; Derijks, L. J. J.	Clinical Rheumatology	Therapeutic drug monitoring (TDM) as a tool in the switch from infliximab innovator to biosimilar in rheumatic patients: results of a 12-month observational prospective cohort study	Population
E. M. H.; Boekema Schmitz, P. J.; Straathof, J. W. A.; van Renswouw, D. C.; Brunsveld, L.; Scharnhorst, V.; van de Poll, M. E. C.; Broeren, M. A. C.; Derijks, L. J. J.	Alimentary Pharmacology & Therapeutics	Switching from infliximab innovator to biosimilar in patients with inflammatory bowel disease: a 12-month multicentre observational prospective cohort study	Population
E. M. H.; van de Kerkhof Schmitz, D.; Hamann, D.; van Dongen, J. L. J.; Kuijper, P. H. M.; Brunsveld, L.; Scharnhorst, V.; Broeren, M. A. C.	Clinical Chemistry and Laboratory Medicine	Therapeutic drug monitoring of infliximab: performance evaluation of three commercial ELISA kits	Design
T.; Keller Schuster, E.; Krauchi, S.; Bantleon, F.; Weber, J.; Schneider, M.	Journal of Crohn's and Colitis	Performance of the BUHLMANN Quantum Blue Infliximab point-of-care assay dedicated	Design

Authors	Source	Title	Reasons for exclusion
		for therapeutic drug monitoring of serum infliximab trough levels	
P.; Corallini Secchiero, F.; Castellino, G.; Bortoluzzi, A.; Caruso, L.; Bugatti, S.; Bosco, R.; Montecucco, M.; Trotta, F.	Journal of Rheumatology	Baseline serum concentrations of TRAIL in early rheumatoid arthritis: Relationship with response to disease-modifying antirheumatic drugs	Design
T.; Cildag Senturk, S.; Akdam, I.; Gultekin, B.	International Journal of Rheumatic Diseases	Anti-TNF induced autoimmunity	Design
T.; Cildat Senturk, S.; Akdam, I.; Gultekin, B.	Clinical and Experimental Rheumatology	Anti-TNF induced autoimmunity	Design
J.; Hamze Sigaux, M.; Daien, C.; Morel, J.; Krzysiek, R.; Pallardy, M.; Maillere, B.; Mariette, X.; Miceli-Richard, C.	Annals of the Rheumatic Diseases	The lack of antidrug antibodies among patients treated with tocilizumab: A clue to good efficacy profiles when used as monotherapy?	Intervention
J.; Hamze Sigaux, M.; Daien, C.; Morel, J.; Krzysiek, R.; Pallardy, M.; Maillere, B.; Mariette, X.; Miceli-Richard, C.	Joint, Bone, Spine: Revue du Rhumatisme	Immunogenicity of tocilizumab in patients with rheumatoid arthritis	Intervention
F.; Arlestig Siljehult, L.; Eriksson, C.; Rantapaa-Dahlqvist, S.	Scandinavian Journal of Rheumatology	Concentrations of infliximab and anti-drug antibodies in relation to clinical response in patients with rheumatoid arthritis	Population
J. S.; Mostafa Smolen, N.; Huang, X.; Noertersheuser, P.; Klunder, B.; Chen, K.; Kalabic, J.; Sainsbury, I.; Oerlemans, R.; Florentinus, S.; Burmester, G. R.	Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	The value of adalimumab trough levels and clinical assessments in predicting clinical response in patients with established rheumatoid arthritis and an inadequate response to methotrexate	Population
D.; Nguyen Sorrentino, V.; Henderson, C.; Bankole, A.	Inflammatory Bowel Diseases	Therapeutic Drug Monitoring and Clinical Outcomes in Immune Mediated Diseases: The Missing Link	Design

Authors	Source	Title	Reasons for exclusion
F. R.; Valesini Spinelli, G.	Clinical and Experimental Rheumatology	Immunogenicity of anti-tumour necrosis factor drugs in rheumatic diseases	Design
M.; Samasca Spirchez, G.; Bolba, C.; Miu, N.	Pediatric Rheumatology. Conference: 18th Pediatric Rheumatology European Society, PReS Congress. Bruges Belgium. Conference Publication:	Serum tumor necrosis factor alpha increased during remission with Etanercept	Population
E. W.; Wagner St Clair, C. L.; Fasanmade, A. A.; Wang, B.; Schaible, T.; Kavanaugh, A.; Keystone, E. C.	Arthritis & Rheumatism	The relationship of serum infliximab concentrations to clinical improvement in rheumatoid arthritis: results from ATTRACT, a multicenter, randomized, double-blind, placebo-controlled trial	Population
L. K.; Barclay Stamp, M.	Rheumatology	Therapeutic drug monitoring in rheumatic diseases: utile or futile?	Design
V.; Balsa Strand, A.; Al-Saleh, J.; Barile-Fabris, L.; Horiuchi, T.; Takeuchi, T.; Lula, S.; Hawes, C.; Kola, B.; Marshall, L.	Biodrugs	Immunogenicity of Biologics in Chronic Inflammatory Diseases: A Systematic Review	Design
K.; Wessels Stubenrauch, U.; Birnboeck, H.; Ramirez, F.; Jahreis, A.; Schleypen, J.	Clinical Therapeutics	Subset analysis of patients experiencing clinical events of a potentially immunogenic nature in the pivotal clinical trials of tocilizumab for rheumatoid arthritis: Evaluation of an antidrug antibody ELISA using clinical adverse event-driven immunogenicity testing	Intervention
M.; Geborek Svenson, P.; Saxne, T.; Bendtzen, K.	Rheumatology	Monitoring patients treated with anti-TNF-alpha biopharmaceuticals: assessing serum infliximab and anti-infliximab antibodies	Design
T.; Miyasaka Takeuchi, N.; Inoue, K.; Abe, T.; Koike, T.	Modern Rheumatology	Impact of trough serum level on radiographic and clinical response to infliximab plus methotrexate in patients with rheumatoid arthritis: Results from the RISING study	Population

Authors	Source	Title	Reasons for exclusion
T.; Miyasaka Takeuchi, N.; Tatsuki, Y.; Yano, T.; Yoshinari, T.; Abe, T.; Koike, T.	Annals of the Rheumatic Diseases	Baseline tumour necrosis factor alpha levels predict the necessity for dose escalation of infliximab therapy in patients with rheumatoid arthritis	Design
T.; Miyasaka Takeuchi, N.; Tatsuki, Y.; Yano, T.; Yoshinari, T.; Abe, T.; Koike, T.	Annals of the Rheumatic Diseases	Inhibition of plasma IL-6 in addition to maintenance of an efficacious trough level of infliximab associated with clinical remission in patients with rheumatoid arthritis: analysis of the RISING Study	Population
T.; Miyasaka Takeuchi, N.; Inui, T.; Yano, T.; Yoshinari, T.; Abe, T.; Koike, T.	Annals of the Rheumatic Diseases	Both high titer of RF/ACPA at baseline is closely linked with high level of baseline plasma TNF level which resulted in low drug level and low clinical response in infliximab treatment in RA Patients: Post-hoc analysis of a double-blind clinical study (rising study)	Design
T.; Miyasaka Takeuchi, N.; Inui, T.; Yano, T.; Yoshinari, T.; Abe, T.; Koike, T.	Arthritis Research and Therapy	High titers of both rheumatoid factor and anti-CCP antibodies at baseline in patients with rheumatoid arthritis are associated with increased circulating baseline TNF level, low drug levels, and reduced clinical responses: A post hoc analysis of the RISING study	Population
T.; Tatsuki Takeuchi, Y.; Yano, T.; Yoshinari, T.; Miyasaka, N.; Abe, T.; Koike, T.	Arthritis and Rheumatism. Conference: Annual Scientific Meeting of the American College of Rheumatology and Association of Rheumatology Health Professionals	Clinical efficacy of infliximab is maximized when both circulating TNF and IL-6 are suppressed in the treatment of rheumatoid arthritisresults from the rising study	Population
R.; Guiducci Terenzi, S.; Nacci, F.; Romano, E.; Manetti, M.; Peruzzi, F.; Bruni, C.; Bartoli, F.; Matucci-Cerinic, M.	Annals of the Rheumatic Disease. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Soluble FAS/FASL levels in rheumatoid arthritis patients treated with infliximab and adalimumab	Population

Authors	Source	Title	Reasons for exclusion
J.; Chamaida Teresa, P. R.; Ana, M. F.; Victoria, N. C.; Theo, R.; Annick, V.; Karien, B.; Eva-Maria, O.; Cristina, D.; Alejandro, V.; Diana, P.; Laura, N.; Maria-Gema, B.; Alejandro, B.; Dora, P. S.	The open rheumatology journal	Predictive Value of Serum Infliximab Levels at Induction Phase in Rheumatoid Arthritis Patients	Population
D.; Ducourau Ternant, E.; Fuzibet, P.; Vignault, C.; Watier, H.; Lequerre, T.; Le Loet, X.; Vittecoq, O.; Goupille, P.; Mulleman, D.; Paintaud, G.	British Journal of Clinical Pharmacology	Pharmacokinetics and concentration-effect relationship of adalimumab in rheumatoid arthritis	Design
D.; Fuzibet Ternant, P.; Ducourau, E.; Vittecoq, O.; Lequerre, T.; Goupille, P.; Mulleman, D.; Paintaud, G.	Fundamental and Clinical Pharmacology	Adalimumab pharmacokinetics and concentration-effect relationship in rheumatoid arthritis	Design
S. S.; Borazan Thomas, N.; Barroso, N.; Duan, L.; Taroumian, S.; Kretzmann, B.; Bardales, R.; Elashoff, D.; Vangala, S.; Furst, D. E.	BioDrugs	Comparative Immunogenicity of TNF Inhibitors: Impact on Clinical Efficacy and Tolerability in the Management of Autoimmune Diseases. A Systematic Review and Meta-Analysis	Design
X.; Su Tian, Y.; He, D.; Zhang, Z.; Zhang, F.	International Journal of Rheumatic Diseases	A prospective open-label study comparing immunogenicity and clinical efficacy of etanercept and infliximab in Chinese patients with RA or AS	Population
C.; Plasencia Tornero, C.; Pascual, D.; Jurado, T.; Monjo, I.; Paredes, M. B.; Moral, E.; Pieren, A.; Nuno, L.; Bonilla, G.; Peitedo, D.; Mola, E. M.; Balsa, A.	Annals of the Rheumatic Diseases	Tapering strategy in patients with rheumatoid arthritis receiving tocilizumab	Intervention
C.; Plasencia Tornero Marin, C.; Pascual Salcedo, D.; Jurado, T.; Paredes, M. B.; Monjo, I.; Moral, E.; Pieren, A.; Bonilla Hernan, G.;	Annals of the Rheumatic Diseases	Tocilizumab serum trough levels correlate with clinical activity in rheumatoid arthritis	Intervention

Authors	Source	Title	Reasons for exclusion
Peiteado, D.; Bogas, P.; Nuno, L.; Villalba Yllan, A.; Martin Mola, E.; Balsa Criado, A.			
L.; Van Den Bemt Tweehuysen, B. J. F.; Van Ingen, I. L.; De Jong, A. J. L.; Van Der Laan, W. H.; Van Den Hoogen, F. H. J.; Den Broeder, A. A.	Arthritis and Rheumatology	Clinical and immunogenicity outcomes after switching treatment from innovator infliximab to biosimilar infliximab in rheumatic diseases in daily clinical practice	Population
L.; Van Den Ende Tweehuysen, C.; Beeren, F.; Been, E.; Van Den Hoogen, F.; Den Broeder, A.	Annals of the Rheumatic Diseases	Prediction of successful dose reduction or discontinuation of biologics in patients with rheumatoid arthritis: A systematic review	Design
L.; Van Den Ende Tweehuysen, C. H.; Beeren, F. M. M.; Been, E. M. J.; Van Den Hoogen, F. H. J.; Den Broeder, A. A.	Arthritis and Rheumatology	No strong evidence supporting predictors for successful dose reduction or discontinuation of a biologic in rheumatoid arthritis: A systematic review	Design
L.; van den Ende Tweehuysen, C. H.; Beeren, F. M.; Been, E. M.; van den Hoogen, F. H.; den Broeder, A. A.	Arthritis & Rheumatology	Little Evidence for Usefulness of Biomarkers for Predicting Successful Dose Reduction or Discontinuation of a Biologic Agent in Rheumatoid Arthritis: A Systematic Review	Design
L.; van den Ende Tweehuysen, C. H.; Beeren, F. M. M.; Been, E. M. J.; van den Hoogen, F. H. J.; den Broeder, A. A.	Arthritis and Rheumatology	Little Evidence for Usefulness of Biomarkers for Predicting Successful Dose Reduction or Discontinuation of a Biologic Agent in Rheumatoid Arthritis: A Systematic Review	Design
L.; Hernandez-Florez Valor, D.; de la Torre, I.; del Rio, T.; Nieto, J. C.; Gonzalez, C.; Lopez-Longo, F. J.; Monteagudo, I.; Llinares, F.; Rosas, J.; Garrido, J.; Naredo, E.; Carreno, L.	Clinical and Experimental Rheumatology	Investigating the link between disease activity and infliximab serum levels in rheumatoid arthritis patients	Design
L.; Hernandez-Florez Valor, D.; de la Torre, I.; Llinares, F.; Rosas, J.; Yague, J.; Garrido, J.; Naredo, E.	Clinical & Experimental Rheumatology	Agreement in assessment of infliximab and adalimumab levels in rheumatoid arthritis: interlaboratory and interassay comparison	Design

Authors	Source	Title	Reasons for exclusion
L.; Hernandez Florez Valor, D.; De La Torre, I.; Llinares, F.; Rosas, J.; Yaque, J.; Naredo, E.; Gonzalez, C.; Lopez-Longo, J.; Monteagudo, I.; Montoro, M.; Carreno Perez, L.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Infliximab and adalimumab levels and antidrug antibodies detection in patients with rheumatoid arthritis (RA): An interlaboratory comparison using a commercial Elisa assay	Design
J. S.; Koch van Bezooijen, B. C.; van Doorn, M. B.; Prens, E. P.; van Gelder, T.; Schreurs, M. W.	Therapeutic Drug Monitoring	Comparison of Three Assays to Quantify Infliximab, Adalimumab, and Etanercept Serum Concentrations	Design
B.; Den Broeder Van den Bemt, A. A.; Wolbink, G. J.; Hekster, Y. A.; Van Riel, Plcm; Benraad, B.; Van den Hoogen, F. H. J.	Pharmacy World & Science	Predictive value of infliximab serum trough levels for response in patients with rheumatoid arthritis	Population
B. J.; den Broeder van den Bemt, A. A.; Snijders, G. F.; Hekster, Y. A.; van Riel, P. L.; Benraad, B.; Wolbink, G. J.; van den Hoogen, F. H.	Annals of the Rheumatic Diseases	Sustained effect after lowering high-dose infliximab in patients with rheumatoid arthritis: a prospective dose titration study	Population
B. J.; Den Broeder Van Den Bemt, A. A.; Wolbink, G.; Hekster, Y. A.; Van Riel, P. L.; Benraad, B.; Van Den Hoogen, F. H.	BMC Musculoskeletal Disorders	Anti-infliximab antibodies are already detectable in most patients with rheumatoid arthritis halfway through an infusioncycle: An open-label pharmacokinetic cohort study	Population
B. J. F.; den Broeder van den Bemt, A. A.; Wolbink, G. J.; van den Maas, A.; Hekster, Y. A.; van Riel, Plcm; Benraad, H. B.; van den Hoogen, F. H. J.	British Journal of Clinical Pharmacology	The combined use of disease activity and infliximab serum trough concentrations for early prediction of (non-)response to infliximab in rheumatoid arthritis	Population
B. J. F.; Den Broeder Van Den Bemt, A. A.	Pharmaceutisch Weekblad	Therapeutic drug monitoring of tumour necrosis factor inhibitors in rheumatoid arthritis. [Dutch]	Design
C. J.; Voskuyl van der Laken, A. E.; Roos, J. C.; Stigter van Walsum, M.;	Annals of the Rheumatic Diseases	Imaging and serum analysis of immune complex formation of radiolabelled infliximab and anti-infliximab in responders and non-	Design

Authors	Source	Title	Reasons for exclusion
de Groot, E. R.; Wolbink, G.; Dijkmans, B. A.; Aarden, L. A.		responders to therapy for rheumatoid arthritis.[Reprint in Ned Tijdschr Geneeskd. 2008 Jul 26;152(30):1672-7; PMID: 18714521]	
M. P. M.; Batstra Van Der Linden, M. R.; Bakker-Jonges, L. E.; Detert, J.; Bastian, H.; Scherer, H. U.; Toes, R. E. M.; Burmester, G. R.; Mjaavatten, M. D.; Kvien, T. K.; Huizinga, T. W. J.; Van Der Helm-Van Mil, A. H. M.	Arthritis and Rheumatism	Toward a data-driven evaluation of the 2010 American College of Rheumatology/European League Against Rheumatism criteria for rheumatoid arthritis: Is it sensible to look at levels of rheumatoid factor?	Population
A.; Den Broeder Van Der Maas, A. A.; Wolbink, G. J.; Van Den Hoogen, F. H. J.; Van Riel, P. L. C. M.; Van Den Bemt, B. J. F.	Arthritis and Rheumatism	Prevalence and persistence of low infliximab serum trough levels in RA patients with low disease activity in daily clinical practice	Design
A.; van den Bemt van der Maas, B. J.; Wolbink, G.; van den Hoogen, F. H.; van Riel, P. L.; den Broeder, A. A.	BMC Musculoskeletal Disorders	Low infliximab serum trough levels and anti-infliximab antibodies are prevalent in rheumatoid arthritis patients treated with infliximab in daily clinical practice: results of an observational cohort study	Population
A.; Van Den Bemt Van Der Maas, B.; Van Den Hoogen, F.; Van Riel, P.; Den Broeder, A.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Can baseline (anti-)infliximab serum trough levels predict successful down-titration or discontinuation of infliximab in rheumatoid arthritis patients with long term low disease activity?	Design
A.; Van Den Bemt Van Der Maas, B.; Van Der Hoogen, F.; Van Riel, P.; Den Broeder, A.	International Journal of Clinical Pharmacy	Baseline (anti-)infliximab serum trough levels do not predict successful down-titration or cessation of infliximab in Rheumatoid Arthritis patients with long term low disease activity	Design
Y.; Te Velthuis Van Hensbergen, H.	Annals of the Rheumatic Diseases	Ready to use CE-IVD smart ELISA kits from sanquin for infliximab and adalimumab levels correlate with the golden standard and can be	Design

Authors	Source	Title	Reasons for exclusion
		used for optimisation of personalised treatment in RA patients	
N.; Bouman van Herwaarden, C. A.; van der Maas, A.; van Vollenhoven, R. F.; Bijlsma, J. W.; van den Hoogen, F. H.; den Broeder, A. A.; van den Bemt, B. J.	Annals of the Rheumatic Diseases	Adalimumab and etanercept serum (anti)drug levels are not predictive for successful dose reduction or discontinuation in rheumatoid arthritis	Outcome
N.; Van Den Bemt Van Herwaarden, B. J. F.; Wientjes, M. H. M.; Kramers, C.; Den Broeder, A. A.	Expert Opinion On Drug Metabolism & Toxicology	Clinical utility of therapeutic drug monitoring in biological disease modifying anti-rheumatic drug treatment of rheumatic disorders: a systematic narrative review	Design
P. A.; Bartelds van Schouwenburg, G. M.; Hart, M. H.; Aarden, L.; Wolbink, G. J.; Wouters, D.	Journal of Immunological Methods	A novel method for the detection of antibodies to adalimumab in the presence of drug reveals "hidden" immunogenicity in rheumatoid arthritis patients	Design
P. A.; Krieckaert van Schouwenburg, C. L.; Rispens, T.; Aarden, L.; Wolbink, G. J.; Wouters, D.	Annals of the Rheumatic Diseases	Long-term measurement of anti-adalimumab using pH-shift-anti-idiotype antigen binding test shows predictive value and transient antibody formation	Design
P. A.; Rispens van Schouwenburg, T.; Wolbink, G. J.	Nature Reviews Rheumatology	Immunogenicity of anti-TNF biologic therapies for rheumatoid arthritis	Design
T.; Lu Van Stappen, J.; Geukens, N.; Spasic, D.; Delport, F.; Zali, N.; Kolmel, Y.; Rameil, S.; Lammertyn, J.; Vande Casteele, N.; Gils, A.	United European Gastroenterology Journal	Point-of-care assays for rapid quantification of infliximab	Design
T.; Vande Casteele Van Stappen, N.; Van Assche, G.; Ferrante, M.; Vermeire, S.; Gils, A.	Gut	Clinical relevance of detecting anti-infliximab antibodies with a drug-tolerant assay: post hoc analysis of the TAXIT trial	Population

Authors	Source	Title	Reasons for exclusion
M.; Guillou Verdet, C.; Potier, M. L.; Hiron, M.; Jouen, F.; Boyer, O.; Lequerre, T.; Vittecoq, O.	Arthritis and Rheumatism	Immunogenicity of infliximab is related to reduction of frequency of infliximab administration in rheumatoid arthritis and spondyloarthritis patients	Population
M.; Guillou Verdet, C.; Golinski, M. L.; Hiron, M.; Jouen, F.; Boyer, O.; Lequerre, T.; Vittecoq, O.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Prolonging between-infusions interval is associated with positivity to anti-infliximab antibodies in rheumatoid arthritis and spondyloarthritis patients	Population
A.; Plasencia Villalba, C.; Peiteado, D.; Nuno, L.; Bonilla, G.; Lojo, L.; Pascual, D.; Del Moral, R.; Lopez Casla, M. T.; Balsa, A.; Martin Mola, E.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	Influence of immunogenicity of anti-TNF therapy in RA patients with a long-term treatment with infliximab or adalimumab	Population
A.; Navarro Compan Villalba Yllan, M. V.; Plasencia Rodriguez, C.; Peiteado Lopez, D.; Bonilla Hernan, G.; Nuno Nuno, L.; Pascual-Salcedo, D.; Olariaga, E.; Balsa Criado, A.; Martin Mola, E.	Annals of the Rheumatic Diseases	Influence of body mass index (BMI) on serum levels of infliximab in patients with rheumatoid arthritis (RA)	Design
F. B.; Morand Vincent, E. F.; Murphy, K.; Mackay, F.; Mariette, X.; Marcelli, C.	Annals of the Rheumatic Diseases	Antidrug antibodies (ADAb) to tumour necrosis factor (TNF)-specific neutralising agents in chronic inflammatory diseases: A real issue, a clinical perspective	Design
F. B.; Pavy Vincent, S.; Krzysiek, R.; Lequerre, T.; Sellam, J.; Taoufik, Y.; Mariette, X.; Miceli-Richard, C.	Joint Bone Spine	Effect of serum anti-tumour necrosis factor (TNF) drug trough concentrations and antidrug antibodies (ADAb) to further anti-TNF short-term effectiveness after switching in rheumatoid arthritis and axial spondyloarthritis	Design
E.; Hebing Vogelzang, R.; Nurmohamed, M.; L'Ami, M.; Krieckaert, C.; Wolbink, G.	Annals of the Rheumatic Diseases	Assessing adherence of RA patients treated with etanercept using etanercept serum trough concentrations and patient self-report	Design

Authors	Source	Title	Reasons for exclusion
E. H.; Pouw Vogelzang, M. F.; Nurmohamed, M.; Kneepkens, E. L.; Rispens, T.; Wolbink, G. J.; Krieckaert, C. L. M.	Annals of the Rheumatic Diseases	Adalimumab trough concentrations in patients with rheumatoid arthritis and psoriatic arthritis treated with concomitant disease-modifying antirheumatic drugs	Population
E.; Kneepkens Vogelzang, E.; Nurmohamed, M.; Van Kuijk, A.; Rispens, T.; Wolbink, G.; Krieckaert, C.	Annals of the Rheumatic Diseases. Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR	A diminished clinical response at 28 and 52 weeks of adalimumab treatment in patients with psoriatic arthritis is associated with anti-drug antibodies	Population
J.; Jokiranta Westerlund, T. S.	Scandinavian Journal of Rheumatology	Monitoring of tnf-alpha blockers infliximab and adalimumab bymeasuringtroughlevel concentrations and anti-drug antibodies	Design
G.; Goupille Wolbink, P.; Sandborn, W.; Marotte, H.; Mulleman, D.; Ternant, D.; Paul, S.; De Longueville, M.; Vande Casteele, N.; Zamacona, M.; O'Brien, C.; Kvien, T. K.; Kavanaugh, A. F.	Arthritis and Rheumatology	Association between plasma certolizumab pegol concentration and improvement in disease activity in rheumatoid arthritis and Crohn's disease	Population
G. J.; Aarden Wolbink, L. A.; Dijkmans, B. A. C.	Current Opinion in Rheumatology	Dealing with immunogenicity of biologicals: Assessment and clinical relevance	Design
G. J.; Voskuyl Wolbink, A. E.; Lems, W. F.; de Groot, E.; Nurmohamed, M. T.; Tak, P. P.; Dijkmans, B. A.; Aarden, L.	Annals of the Rheumatic Diseases	Relationship between serum trough infliximab levels, pretreatment C reactive protein levels, and clinical response to infliximab treatment in patients with rheumatoid arthritis	Population
G. J.; Vis Wolbink, M.; Lems, W.; Voskuyl, A. E.; de Groot, E.; Nurmohamed, M. T.; Stapel, S.; Tak, P. P.; Aarden, L.; Dijkmans, B.	Arthritis & Rheumatism	Development of antiinfliximab antibodies and relationship to clinical response in patients with rheumatoid arthritis	Population
P.; Bowling Wong, A.; Ford, E.; Freeman, D.; Bagga, H.; Langguth, D.	Internal Medicine Journal	Serum trough levels of adalimumab and infliximab inversely correlate with disease activity in patients with inflammatory arthritis	Population

Authors	Source	Title	Reasons for exclusion
C.; Wang Wu, S.; Xian, P.; Yang, L.; Chen, Y.; Mo, X.	BioMed Research International	Effect of Anti-TNF Antibodies on Clinical Response in Rheumatoid Arthritis Patients: A Meta-Analysis	Design
M.; Becher Zanker, G.; Arbach, O.; Maurer, M.; Stuhlmuller, B.; Schafer, A.; Strohner, P.; Brand, J.	Clinical & Experimental Rheumatology	Improved adalimumab dose decision with comprehensive diagnostics data	Population
M.; Becher Zanker, G.; Arbach, O.; Maurer, M.; Stuhlmuller, B.; Schafer, A.; Strohner, P.; Brand, J.	Clinical and Experimental Rheumatology	Improved adalimumab dose decision with comprehensive diagnostics data	Population
P.; Jeanfavre Zufferey, M. F. D.; Dumusc, A.; Benaim, C.; Perreau, M.; So, A. K.	Arthritis and Rheumatology	Is it possible to predict which patients treated with biologic agents for rheumatic diseases will develop anti-drug antibodies ?	Design
Jani M, Chinoy H, Warren RB, Griffiths CEM, Plant D, Fu B, Morgan AW, Wilson AW, Isaacs JD, Hyrich KL, Barton AB on behalf of BRAGGSS.	Arthritis and Rheumatology	Clinical utility of random anti-tumour necrosis factor drug testing and measurement of anti-drug antibodies on long-term treatment response in rheumatoid arthritis	Population
Jani, M, Isaacs, J. D., Morgan, A. W., Wilson, A. G., Plant, D., Hyrich, K. Chinoy, H, Barton, A.	Annals of the Rheumatic Diseases	High frequency of antidrug antibodies and association of random drug levels with efficacy in certolizumab pegol-treated patients with rheumatoid arthritis: results from the BRAGGSS cohort	Population
Jani, M, Isaacs, J. D., Morgan, A. W., Wilson, A. G., Plant, D., Hyrich, K, Chinoy H, Barton, A	Rheumatology	Detection of anti-drug antibodies using a bridging ELISA compared with radioimmunoassay in adalimumab-treated rheumatoid arthritis patients with random drug levels	Population
M. Jani, W.G. Dixon, M. Lunt, D. De Cock, J.D. Isaacs, A.W. Morgan, A.G.	Ann Rheum Dis	The association of biologic drug-levels with infection risk: results from the british society for	Population

Authors	Source	Title	Reasons for exclusion
Wilson, D. Plant, K. Watson, A. Barton, K. Hyrich		rheumatology biologics register for rheumatoid arthritis Ann Rheum	
Jani M, Dixon WG, Lunt M, De Cock D, Isaacs J, Morgan A, Watson K, Wilson AG, Barton A, Hyrich KL	Arthritis Rheumatol	The Association of Biologic Drug-Levels with Infection Risk: Results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis	Population
l'Ami MJ, Krieckaert CL, Nurmohamed MT, van Vollenhoven RF, Rispens T, Boers M, Wolbink GJ	Ann Rheum Dis	Successful reduction of overexposure in patients with rheumatoid arthritis with high serum adalimumab concentrations: an open-label, non-inferiority, randomised clinical trial	Population

Appendix 3. Quality Assessment

A3.1 (Part I): Quality assessment on the basis of specific outcomes

Table below presents the risk of bias assessment on the basis of specific outcomes: clinical disease activity (disease flare, remission, and change in disease activity), proportion of patients receiving dose tapering, health-related quality of life (HQoL), discontinuation of treatment and treatment dose-related outcomes. For each specific outcome, the following bias domains were assessed: bias due to confounding, bias in selection of participants into the study, bias in measurement of interventions, bias due to departures from intended interventions, bias due to missing data, bias in taking measurements, and bias in selection of the reported result.

In terms of outcome-specific assessments, both the treatment dose-related outcome and the outcome of clinical activity (disease flare, remission and change in disease activity) were judged to be at moderate risk of bias, given that there was moderate risk in the domain of bias due to confounding. For both outcomes, there were low to moderate risks of biases for the remaining bias domains: bias in selection of participants into the study, bias in measurement of interventions, bias in taking measurements, and bias in selection of the reported results.

The outcome of discontinued treatment was judged to be at moderate of bias because there was moderate risk in the domain of bias due to missing data. For this outcome, there were low to moderate risks of biases for the remaining bias domains: bias in selection of participants into the study, bias in measurement of interventions, bias due to departures from intended interventions, bias in taking measurements, and bias in selection of the reported results.

Regarding outcomes of health related quality of life (HRQoL) and proportion of patients receiving dose tapering, both outcomes were judged to be at moderate risk of bias because there was moderate risk of bias for two bias domains (bias in taking measurements and bias due to confounding). For both outcomes, there was low risk of bias for the remaining bias domains: bias in selection of participants into the study, bias in measurement of interventions, and bias in selection of the reported results.

Risk of bias in outcome-specific assessments

Domain	Clinical activity (disease flare, remission, change in disease activity)	Proportion tapered	Health related quality of life	Discontinued treatment	Dose-related outcomes
Bias due to confounding	Moderate - serious	Moderate	Serious	NA*	Moderate
Bias in selection of participants into the study	Low - moderate	Low	Low	Moderate	Low – moderate
Bias in measurement of interventions	Low - moderate	Low	Low	Low	Low - moderate
Bias due to departures from intended interventions	NI	NI	NI	Moderate	NI
Bias due to missing data	Low to serious	NI	Serious	Moderate	NI
Bias in taking measurements	Moderate	Moderate	Moderate	Moderate	Moderate
Bias in selection of the reported result	Low	Low	Low	Low	Low
Overall risk of bias	Serious	Moderate	Serious	Moderate	Moderate

Key: Risk of bias judgement: low/moderate/serious/critical/NI. NI: no information; NA*: not applicable because of lack of comparator group; HQoL: health-related quality of life

A3.2 (Part II): Quality assessment of individual studies

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **ARANGO 2017**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent)
Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized√ / Cluster randomized / Matched (e.g. cross-over)
Participants	Adult patient treated with Adalimumab (40mg sc) who remained clinically stable for at least 6 months
Experimental intervention	Adjustment of ADL frequency (tapering) plus therapeutic drug monitoring (TDM) data to revealed to physicians
Comparator	Adjustment of ADL frequency (tapering), physicians blinded to TDM data

Is your aim for this study…?

□ to assess the effect of *assignment to* intervention
√ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Proportion of patients tapered (benefit), rate of flare (harm)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, **figure** or paragraph) that uniquely defines the result being assessed.

Proportion tapered; 34.6% (CG), 35.7% (IG)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)		No	Yes	Expected to favour control group (28.6% IG had LDA vs. 17.3% of CG)
Time of assessment for response		No	No information	No information but likely to be unimportant. Measurement believed to be done at similar time points (at 8 scheduled visits over 18 months)
Serum Adalimumab levels		No	Yes	NA – serum ADL levels 5.76mg/L in the CG and 5.04mg/L in IG.
Serum anti-Adalimumab antibody levels		No	No information	No information

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
Methotrexate and other DMARDs	No	Favour experimental / Favour comparator / No information√
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Yes, differential baseline LDA rates and no information on co-intervention	Y / PY √ / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N √ / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N √ / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA√ / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA√ / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N√ / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4		Y / PY / PN / N√ / NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		NA √ / Y / PY / PN / N / NI
Risk of bias judgement		Low √ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?		Y √ / PY / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y√ / PY / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA√ / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		Y / PY / PN / N / NI√
4.4. Was the intervention implemented successfully for most participants?		Y √ / PY / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY√ / PN / N / NI
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?		Y / PY / PN / N / NI√
5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN / N / NI√

5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		NA√ / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN √ / N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y √ / PY / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?		Y / PY / PN / N / NI√
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N√ / NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **CHEN 2016**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent) Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	Patients in remission/low disease activity
Experimental intervention	Dose halving of Adalimumab plus therapeutic drug monitoring of serum adalimumab and anti-drug antibody (Sandwich and bridging ELISA)
Comparator	None

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Persistent remission/LDA (beneficial) and disease flare (harmful)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Persistent remission/LDA = 23 + 24 = 47/64 (73.4%); disease flare = 15/64 (23.4%)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)				
Time of assessment for response				
Serum Adalimumab levels				
Serum anti-Adalimumab antibody levels				

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
Methotrexate		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered		Y / PY / PN / **N**√
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	*'We enrolled 64 initially biologic-naïve patients who fulfilled the 1987 ACR criteria for RA and had achieved remission or LDA after receiving ADA full-dose therapy……'*	Y / PY / PN / N √/ NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Adalimumab dose-halving (40mg monthly) and a concomitant stable dose of methotrexate	Y√ / PY / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y √/ PY / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Not relevant	Y / PY / PN / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	Not relevant	Y / PY / PN / N / NI
4.4. Was the intervention implemented successfully for most participants?		Y / PY√ / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY √/ PN / N / NI
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		moderate
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	*'after 24 weeks of dose-halving, persistent remission was observed in 23 patients, remission turned LDA in 2, persistent LDA in 24 and disease flare in 15 patients…….'*	Y√ / PY / PN / N / NI

5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN / N√ / NI
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N√ / NI
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		NA√ / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN / N√ / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y / PY √/ PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?	Not applicable	Y / PY / PN / N / NI Not applicable
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN√ / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N √/ NI
7.3 ... different *subgroups*?		Y / PY / PN / N √/ NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Superseded – see Erratum

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **GOROSTIZA 2016**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent)
Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized√ / Cluster randomized / Matched (e.g. cross-over)
Participants	Patient treated with adalimumab (40mg sc) who remained clinically stable for at least six months
Experimental intervention	Biological monitoring data (BMD) were released to physicians
Comparator	Physicians were blinded to BMD data

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Proportion remaining in remission (benefit)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Promotion remaining in remission = 69.6% (CG), 76.1% (IG)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)		No	Yes	Expected to favour control group (26.6% IG had LDA vs. 16.7% of CG)
Time of assessment for response		No	No information	No information but likely to be unimportant. Measurement believed to be done at similar time points (at 8 scheduled visits over 18 months)
Serum Adalimumab levels		No	Yes	NA – serum ADL levels 5.5mg/L in the CG and 5.3mg/L in IG.
Serum anti-Adalimumab antibody levels		No	No information	No information

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
Methotrexate and other DMARDs	No	Favour experimental / Favour comparator / No information√
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Yes, differential baseline LDA rates and no information on co-intervention	Y / PY √/ PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N√ / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N √/ NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA √/ Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA√ / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N√ / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4		Y / PY / PN / N√ / NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	BDM data were released only to IG	Y√ / PY / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y√ / PY / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		√NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		Y / PY / PN / N / NI√
4.4. Was the intervention implemented successfully for most participants?		Y √/ PY / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY√ / PN / N / NI
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?		Y / PY / PN / N / NI√
5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN / N / NI√

5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		NA√ / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN √/ N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y√ / PY / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?		Y / PY / PN / N / NI√
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N√ / NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **INCIARTE-MUNDO 2016**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent)
Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	Patients in clinical remission or low disease activity
Experimental intervention	Treated with adalimumab, etanercept and infliximab; drug serum levels measured every 4 months for 1 year
Comparator	None

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Disease flare

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Disease flare = 17.0% among Rheumatoid arthritis patients

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)				NA
Time of assessment for response				NA
Serum Adalimumab levels				NA
Serum anti-Adalimumab antibody levels				NA

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
csDMARD and bDMARD	NA (no comparator group)	Favour experimental / Favour comparator / No information√
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered		Y / PY / PN / N√
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	'……patients in clinical remission (CR) (DAS28-ESR < 2.6) or low disease activity (LDA) (DAS28-ESR < 3.2) in ≥ 2 consecutive visits treated with adalimumab (ADA), etanercept (ETN) or infliximab (IFX) for ≥ 3 months…….'	Y / PY / PN / N√ / NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	'……patients in clinical remission (CR) (DAS28-ESR < 2.6) or low disease activity (LDA) (DAS28-ESR < 3.2) in ≥ 2 consecutive visits treated with adalimumab (ADA), etanercept (ETN) or infliximab (IFX) for ≥ 3 months…….' No control group	Y / PY√ / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y/ PY√ / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Not relevant	Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	Not relevant	Y / PY / PN / N / NI
4.4. Was the intervention implemented successfully for most participants?		Y / PY√ / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY √/ PN / N / NI
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		NA / Y / PY / PN / N / NI
Risk of bias judgement		moderate
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	12 patients (8 RA, 4 psoriasis) with flare was reported as 13% implying the denominator was lower than the total population of 103 (47 RA, 56 PsA)	Y / PY/ PN √/ N / NI
5.2 Were participants excluded due to missing data on intervention status?		Y / PY√ / PN / N / NI
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?	Not relevant (no control group)	√NA / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA / Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN / N√ / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y / PY√ / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?	Not applicable (No control group)	Y / PY / PN / N / NI
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN √/ N / NI
Risk of bias judgement		Low / Moderate √/ Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN√ / N / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN √/ N / NI
7.3 ... different *subgroups*?		Y / PY / PN√ / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate √/ Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **LOPEZ-CASLA 2013**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent) Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	Primary/secondary non-responders
Experimental intervention	Infliximab dose escalation plus therapeutic drug monitoring
Comparator	None

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Treatment discontinuation rate (harmful)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Treatment discontinuation rate = 26 (76.5%)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)	Confounding domain not applicable as there is no comparison group			
Time of assessment for response				
Serum Adalimumab levels				
Serum anti-Adalimumab antibody levels				
(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Methotrexate, other DMARDs, combination or monotherapy	Domain not applicable as there is no comparison group			

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Domain not applicable	Y / PY / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	'Study enrolled 36 RA patients treated since 2000 with ifx at La Paz University Hospital, in whom a Ifx dose increase was implemented due to inefficacy…..'	Y / PY / PN / N √/ NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY / PN √/ N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	'…… In whom a ifx dose increase was implemented due to inefficacy….'	Y / PY √/ PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y / PY√ / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Not relevant	Y / PY / PN / N / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		Not aplicable
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	Not relevant	Y / PY / PN / N / NI
4.4. Was the intervention implemented successfully for most participants?		Y / PY √/ PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY√ / PN / N / NI
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		moderate
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	26 of 36 (baseline denominator) = 72.2%, but this was reported as 76.5% implying two patients were not accounted for in the final analysis	Y / PY / PN√ / N / NI

5.2 Were participants excluded due to missing data on intervention status?		Y / PY/ PN / N / NI√
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		√NA / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA / Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN / N√ / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y√ / PY / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?	Not applicable	Y / PY / PN / N / NI NA√
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low / Moderate √/ Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N √/ NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Superseded – see Erratum

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **PARADES 2015**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent) Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	RA patients in remission or LDA for at least 6 months
Experimental intervention	Infliximab, adalimumab or etanercept dose optimization strategy (OS) (tapering)
Comparator	None

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Flare (harmful)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Rate of flares = 19 (35.2%)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)	Confounding domain not applicable as there is no comparison group			
Time of assessment for response				
Serum Adalimumab levels				
Serum anti-Adalimumab antibody levels				

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
Methotrexate, other DMARDs, combination or monotherapy	Domain not applicable as there is no comparison group	Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Confounding not applicable	Y / PY / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	'…..a cohort of RA patients in LDA or remission (measured by DAS28 < 3.2 or < 2.6, respectively) for at least 6 months..'	Y / PY / PN / N√ / NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Optimization strategy	Y / PY√ / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y / PY√ / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Not relevant	Y / PY / PN / N / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		Not applicable
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	Not applicable	Y / PY / PN / N / NI
4.4. Was the intervention implemented successfully for most participants?		Y / PY √/ PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY / PN / N / NI√
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		NI
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	Proportion of flares to baseline population is appropriate (19; 35.2%)	√Y / PY / PN / N / NI
5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN√ / N / NI

5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN√ / N / NI
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		√NA / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN√ / N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y / PY√ / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?	Not applicable	Y / PY / PN / N / NI Not applicable
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN √/ N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N√ / NI
7.3 ... different *subgroups*?		Y / PY / PN / N √/ NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **PAREDES 2016**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent) Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	RA patients in remission or LDA for at least 6 months
Experimental intervention	Infliximab, adalimumab or etanercept dose tapering
Comparator	None

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Flare (harmful)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Flare (final visit) = 12 (23%)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)	Confounding domain not applicable as there is no comparison group			
Time of assessment for response				
Serum Adalimumab levels				
Serum anti-Adalimumab antibody levels				

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
Methotrexate, other DMARDs, combination or monotherapy	Domain not applicable as there is no comparison group	Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Confounding not applicable	Y / PY / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	'……all patients were in LDA or remission (DAS28 < 3.2 or < 2.6, respectively) for at least 6 months prior to start of the tapering strategy……'	Y / PY / PN / N√ / NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Tapering strategy	Y / PY √/ PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y / PY√ / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Not relevant	Y / PY / PN / N / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		Not applicable
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	Not applicable	Y / PY / PN / N / NI
4.4. Was the intervention implemented successfully for most participants?		Y / PY √/ PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY / PN / N / NI√
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		NI
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	Paredes et al 2015 (which appears to be a mid-way report for this study) reported baseline population of 54 as opposed to 52 reported in final assessment	Y / PY / PN√ / N / NI

5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN / N / NI√
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		√NA / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA / Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate√ / Serious/ Critical / NI
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN √/ N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y / PY √/ PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?	Not applicable	Y / PY / PN / N / NI Not applicable
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN√ / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N √/ NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious/ Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Superseded – see Erratum

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **PASCUAL-SALCEDO 2013**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent) Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	RA Patients in remission or LDA
Experimental intervention	Down-titration or cessation of Infliximab, adalimumab, etanercept plus therapeutic monitoring period
Comparator	Down-titration or cessation of Infliximab, adalimumab, etanercept, prior to therapeutic monitoring period

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Mean DAS28 score (harmful), weekly mean dose (lower better), interval of administration (higher better)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Mean DAS28 score; 1st period: 2.51±0.85 vs 2nd period: 2.31±0.52

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)		No	No information	No information
Time of assessment for response		No	No information	No information
Serum Adalimumab levels		No	No information	No information
Serum anti-Adalimumab antibody levels		No	No information	No information

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
Methotrexate, other DMARDs, combination or monotherapy	Not done/no information	Favour experimental / Favour comparator / No information√
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered		Y / PY√ / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N√ / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN√ / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		√NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N√ / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN√ / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	'…….a total of 88 patients (43 RA and 45 SpA), treated with three TNF inhibitors ……were included…'	Y / PY / PN / N√ / NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		√NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		√NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY / PN / N / NI√
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?		Y√ / PY / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y / PY / PN / N√ / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low / Moderate √/ Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		Not applicable
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		Y / PY / PN / N / NI√
4.4. Was the intervention implemented successfully for most participants?		Y / PY√ / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY / PN / N / NI√
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		NI
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	Results reported were basically means (SD); difficult to determine	Y / PY / PN / N / NI√
5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN / N / NI√

5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		√NA / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN√ / N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y √/ PY / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?		Y / PY√ / PN / N / NI
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN√ / N / NI
Risk of bias judgement		Low / Moderate √/ Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N √/ NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **ROSAS 2015**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes, inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent) Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	RA patients on clinical remission
Experimental intervention	ADL, ETN dose frequency decrease plus therapeutic drug monitoring
Comparator	None

Is your aim for this study…?

- □ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Clinical remission maintenance (benefit)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Proportion with modified dosing frequency remaining on clinical remission = 87%

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)	Confounding domain not applicable as there is no comparator group			
Time of assessment for response				
Serum Adalimumab levels				
Serum anti-Adalimumab antibody levels				
(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
Methotrexate, other DMARDs, combination or monotherapy		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Not applicable	Y / PY / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	'……patients with rheumatoid arthritis (RA) in clinical remission……clinical remission was defined as sustained DAS28-ESR ≤ 2.6 during 6 consecutive months….'	Y / PY / PN / N / NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	'…adalimumab (ADL) and etanercept (ETN) dose reduction (by decreasing treatment frequency) and drug monitoring in patients……'	Y √/ PY / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y / PY √/ PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		Y / PY / PN√ / N / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		Not applicable
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	Not applicable	Y / PY / PN / N / NI
4.4. Was the intervention implemented successfully for most participants?		Y / PY√ / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY / PN / N / NI√
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		NI
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?		Y / PY / PN / N / NI√

5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN / N / NI√
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		NA / Y / PY / PN / N / NI√
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA / Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN√ / N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y√ / PY / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?	Not applicable	Y / PY / PN / N / NI Not applicable
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN√ / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N √/ NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N√ / NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate √/ Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Superseded – see Erratum

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **SENABRE GALLEGO 2017**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent)
Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized / Cluster randomized / Matched (e.g. cross-over)
Participants	RA patients in clinical remission
Experimental intervention	Adalimumab or Etanercept with extended interval of administration (tapering) plus therapeutic drug monitoring
Comparator	Adalimumab or Etanercept with extended interval of administration, no drug monitoring (comparator not available for Senabre et al)

Is your aim for this study…?

- ☐ to assess the effect of *assignment to* intervention
- √ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Worsening of clinical activity (harm)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Patients with worsening of clinical activity = 9/39 (23.1%)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)	Confounding domain not applicable as there is no comparator group			
Time of assessment for response				
Serum Adalimumab levels				
Serum anti-Adalimumab antibody levels				

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
Methotrexate, other DMARDs, combination or monotherapy	Not relevant (no comparator)	Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Not applicable	Y / PY / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N / NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N / NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4	'……RA patients in clinical remission, receiving adalimumab (ADL) or etanercept (ETN) with EIA…….'	Y / PY / PN / N √/ NI

2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		√NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		√NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	'…….patients receiving anti-TNFα therapy with extended interval of administration (EIA)…..'	Y√ / PY / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	Prospective	Y √/ PY / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		Y / PY / PN / N √/ NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		Not applicable
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	Not applicable	Y / PY / PN / N / NI
4.4. Was the intervention implemented successfully for most participants?	'……2 patients never reduced anti-TNFα due to low drug levels……'	Y√ / PY / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?	'……one patient was excluded due to blindness violation…..'	Y / PY √/ PN / N / NI
4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		√NA / Y / PY / PN / N / NI
Risk of bias judgement		Moderate
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?		Y √/ PY / PN / N / NI

5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN√ / N / NI
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		√NA / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA / Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Study appears blinded *'……one patient was excluded due to blindness violation…..'*	Y / PY / PN / N √/ NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y / PY / PN√ / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?		Y / PY / PN / N / NI Not applicable
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN√ / N / NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N √/ NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Superseded – see Erratum

The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tool
(version for cohort-type studies) **UCAR 2017**
Version 19 September 2016

ROBINS-I tool (Stage I): At protocol stage

Specify the review question

Participants	Remission/primary non-responders/secondary non-responders
Experimental intervention	Therapeutic drug monitoring
Comparator	Standard care
Outcomes	13 outcomes; inconclusive results, time to results, dose changes, dose adjustment, treatment switch, discontinuation, changes in disease activity, rate of disease response, relapse and remission, hospitalization, rates of surgical intervention, adverse effects, health-related quality of life

List the confounding domains relevant to all or most studies

From protocol; time of testing, testing method (e.g. reflex vs. concurrent) Others (suggested); drug dose/levels, disease stage at enrollment, time of assessment for response/follow-up, type of drug manipulation (e.g. optimisation or tapering)

List co-interventions that could be different between intervention groups and that could impact on outcomes

Methotrexate, other DMARDs, combination or monotherapy

ROBINS-I tool (Stage II): For each study

Specify a target randomized trial specific to the study

Design	Individually randomized √/ Cluster randomized / Matched (e.g. cross-over)
Participants	Patient treated with adalimumab (40mg sc) who remained clinically stable for at least six months
Experimental intervention	Biological monitoring data (BMD) were released to physicians
Comparator	Physicians were blinded to BMD data

Is your aim for this study…?

□ to assess the effect of *assignment to* intervention
√ to assess the effect of *starting and adhering to* intervention

Specify the outcome

Specify which outcome is being assessed for risk of bias (typically from among those earmarked for the Summary of Findings table). Specify whether this is a proposed benefit or harm of intervention.

Disease flare (harm)

Specify the numerical result being assessed

In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

IRR = 0.7252 (95% CI = 0.49997 to 1.0578)

Preliminary consideration of confounders

Complete a row for each important confounding domain (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as potentially important.

"Important" confounding domains are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention. "Validity" refers to whether the confounding variable or variables fully measure the domain, while "reliability" refers to the precision of the measurement (more measurement error means less reliability).

(i) Confounding domains listed in the review protocol				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information
Disease stage (proportion in remission/LDA)		No	Yes	Expected to favour control group (26.6% IG had LDA vs. 16.7% of CG)
Time of assessment for response		No	No information	No information but likely to be unimportant. Measurement believed to be done at similar time points (at scheduled visits)
Serum Adalimumab levels		No	Yes	NA – serum ADL levels 5.5mg/L in the CG and 5.3mg/L in IG.
Serum anti-Adalimumab antibody levels		No	No information	No information

(ii) Additional confounding domains relevant to the setting of this particular study, or which the study authors identified as important				
Confounding domain	Measured variable(s)	Is there evidence that controlling for this variable was unnecessary?*	Is the confounding domain measured validly and reliably by this variable (or these variables)?	OPTIONAL: Is failure to adjust for this variable (alone) expected to favour the experimental intervention or the comparator?
			Yes / No / No information	Favour experimental / Favour comparator / No information

* In the context of a particular study, variables can be demonstrated not to be confounders and so not included in the analysis: (a) if they are not predictive of the outcome; (b) if they are not predictive of intervention; or (c) because adjustment makes no or minimal difference to the estimated effect of the primary parameter. Note that "no statistically significant association" is not the same as "not predictive".

Preliminary consideration of co-interventions

Complete a row for each important co-intervention (i) listed in the review protocol; and (ii) relevant to the setting of this particular study, or which the study authors identified as important.

"Important" co-interventions are those for which, in the context of this study, adjustment is expected to lead to a clinically important change in the estimated effect of the intervention.

(i) Co-interventions listed in the review protocol		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

(ii) Additional co-interventions relevant to the setting of this particular study, or which the study authors identified as important		
Co-intervention	Is there evidence that controlling for this co-intervention was unnecessary (e.g. because it was not administered)?	Is presence of this co-intervention likely to favour outcomes in the experimental intervention or the comparator
Methotrexate and other DMARDs	No	Favour experimental / Favour comparator / No information √
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information
		Favour experimental / Favour comparator / No information

Risk of bias assessment

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions	Description	Response options
Bias due to confounding		
1.1 Is there potential for confounding of the effect of intervention in this study? **If N/PN to 1.1:** the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Yes, differential baseline LDA rates and no information on co-intervention	Y / PY√ / PN / N
If Y/PY to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow up time according to intervention received? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, go to question 1.3.		NA / Y / PY / PN / N√ / NI
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? **If N/PN**, answer questions relating to baseline confounding (1.4 to 1.6) **If Y/PY**, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		NA / Y / PY / PN / N / NI
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?		NA / Y / PY / PN / N √/ NI
1.5. **If Y/PY to 1.4**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA√ / Y / PY / PN / N / NI
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NA√ / Y / PY / PN / N / NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		NA / Y / PY / PN / N √/ NI
1.8. **If Y/PY to 1.7**: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to confounding?		Favours experimental / Favours comparator / Unpredictable

Bias in selection of participants into the study

2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? **If N/PN to 2.1:** go to 2.4		Y / PY / PN / N√ / NI
2.2. **If Y/PY to 2.1**: Were the post-intervention variables that influenced selection likely to be associated with intervention?		NA / Y / PY / PN / N / NI
2.3 **If Y/PY to 2.2**: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		NA / Y / PY / PN / N / NI
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y / PY√ / PN / N / NI
2.5. **If Y/PY to 2.2 and 2.3, or N/PN to 2.4**: Were adjustment techniques used that are likely to correct for the presence of selection biases?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of participants into the study?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	BDM data were released only to IG	Y √/ PY / PN / N / NI
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y√ / PY / PN / N / NI
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		Y / PY / PN / N √/ NI
Risk of bias judgement		Low √/ Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to classification of interventions?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?		Y / PY / PN / N / NI
4.2. **If Y/PY to 4.1**: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		NA√ / Y / PY / PN / N / NI
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		Y / PY / PN / N / NI√
4.4. Was the intervention implemented successfully for most participants?		Y √/ PY / PN / N / NI
4.5. Did study participants adhere to the assigned intervention regimen?		Y / PY√ / PN / N / NI

4.6. **If N/PN to 4.3, 4.4 or 4.5**: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		NA√ / Y / PY / PN / N / NI
Risk of bias judgement		NI
Optional: What is the predicted direction of bias due to deviations from the intended interventions?		

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?		Y / PY / PN / N / NI√
5.2 Were participants excluded due to missing data on intervention status?		Y / PY / PN / N / NI√
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		Y / PY / PN / N / NI√
5.4 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Are the proportion of participants and reasons for missing data similar across interventions?		NA√ / Y / PY / PN / N / NI
5.5 **If PN/N to 5.1, or Y/PY to 5.2 or 5.3**: Is there evidence that results were robust to the presence of missing data?		NA √/ Y / PY / PN / N / NI
Risk of bias judgement		Low / Moderate / Serious / Critical / NI√
Optional: What is the predicted direction of bias due to missing data?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?		Y / PY / PN√ / N / NI
6.2 Were outcome assessors aware of the intervention received by study participants?		Y √/ PY / PN / N / NI
6.3 Were the methods of outcome assessment comparable across intervention groups?		Y / PY / PN / N / NI√
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		Y / PY / PN / N / NI√
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the predicted direction of bias due to measurement of outcomes?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from... 7.1. ... multiple outcome *measurements* within the outcome domain?		Y / PY / PN / N√ / NI
7.2 ... multiple *analyses* of the intervention-outcome relationship?		Y / PY / PN / N√ / NI
7.3 ... different *subgroups*?		Y / PY / PN / N√ / NI

Risk of bias judgement		Low√ / Moderate / Serious / Critical / NI
Optional: What is the predicted direction of bias due to selection of the reported result?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall bias		
Risk of bias judgement		Low / Moderate√ / Serious / Critical / NI
Optional: What is the overall predicted direction of bias for this outcome?		Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Appendix 4. Search strategy for the additional search

The MEDLINE search strategy was:

1. (anti-TNF* or antiTNF* or (TNF* adj2 (inhibit* or block*))).tw.

2. anti* tumo?r* necrosis* factor*.tw.

3. Tumor Necrosis Factor-alpha/

4. (biologic* adj2 DMARD*).tw.

5. ((antirheumati* or anti rheumati* or anti-rheumati*) adj4 biologic*).tw.

6. ((disease modify* or disease-modify*) adj4 biologic*).tw.

7. exp Antibodies, Monoclonal/

8. anti* drug* antibod*.tw.

9. ADAb.tw.

10. etanercept.tw. or ETANERCEPT/

11. (tnr001 or "tnr 001" or tnr-001 or 185243-69-0).tw.

12. (ETA or ETN).tw.

13. (enbrel or erelzi or benepali or lifmior or brenzys).tw.

14. (anti-etanercept* or antietanercept* or (anti adj3 etanercept*)).tw.

15. adalimumab.tw. or ADALIMUMAB/

16. (d 2e7 or d2e7 or d-2e7 or 331731-18-1).tw.

17. (ADA or ADL or ADM).tw.

18. (humira or amgevita or cyltezo or imraldi or solymbic or hyrimoz or halimatoz).tw.

19. (anti-adalimumab* or antiadalimumab* or (anti adj3 adalimumab*)).tw.

20. infliximab.tw. or INFLIXIMAB/

21. (170277-31-3 or ta650 or ta 650 or ta-650).tw.

22. (INF or IFX).tw.

23. (anti-infliximab* or antiinfliximab* or (anti adj3 infliximab*)).tw.

24. (remicade or inflectra or remsima or flixabi or zessly or renflexis or ixifi).tw.

25. Certolizumab Pegol/ or certolizumab.tw.

26. (cdp870 or cdp 870 or cdp-870 or 428863-50-7 or 1132819-27-2).tw.

27. (CER or CZP).tw.

28. cimzia.tw.

29. (anti-certolizumab* or anticertolizumab* or (anti adj3 certolizumab*)).tw.

30. golimumab.tw.

31. (cnto 148 or cnto148 or cnto-148 or 476181-74-5).tw.

32. (GOL or GLM).tw.

33. simponi.tw.

34. (anti-golimumab* or antigolimumab* or (anti adj3 golimumab*)).tw.

35. (biologic* adj2 agent*).tw.

36. (CT-P13 or CTP13 or CT P13 or SB2 or SB-2 or SB 2 or SB4 or SB-4 or SB 4 or SB-5 or SB5 or SB 5).tw.

37. (biosimilar* or (bio* adj1 similar*)).tw.

38. or/1-37

39. exp Arthritis, Rheumatoid/

40. RA.tw.

41. Rheumarthrit*.tw.

42. ((Rheumatoid* or rheumatic* or inflammat* or idiopathic* or deforman*) adj4 (arthrit* or arthros* or polyarthrit* or factor*)).tw.

43. (Chronic* adj4 (polyarthrit* or poly arthrit* or poly-arthrit* or rheumati*)).tw.

44. ((Inflammat* or pain* or swell* or stiff*) adj4 (joint* or synovial*)).tw.

45. (Beauvais* adj2 disease*).tw.

46. or/39-45

47. Radioimmunoassay/

48. (radioimmuno* or radio immuno* or radio-immuno*).tw.

49. RIA.tw.

50. reporter* gene* assay*.tw.

51. RGA.tw.

52. (semi* fluid* phase* adj3 enzyme* immuno*).tw.

53. EIA.tw.

54. ((homogenous* or homogeneous*) adj1 mobilit* shift* assay*).tw.

55. HMSA.tw.

56. (Biomonitor* or iLite or Euro Diagnostica* or Wieslab or Svar).tw.

57. (ARUP or Q-ETA or EURIA).tw.

58. (Matriks* Biotek* or Shikari*).tw.

59. (Prometheus* or Anser*).tw.

60. or/47-59

61. 38 and 46 and 60

62. randomized controlled trial.pt.

63. controlled clinical trial.pt.

64. randomized.ab.

65. placebo.ab.

66. clinical trials as topic.sh.

67. randomly.ab.

68. trial.ti.

69. or/62-68

70. exp animals/ not humans.sh.

71. 69 not 70

Appendix 5. Exeter biologic clinic recommendations for biologic dose reduction

Patient selection:

- Biologic treatment > 2 years & sustained Low Disease Activity or Clinical Remission (DAS 28 <2.6 +/- USS remission) or BASDAI & Pain VAS <4
- No radiographic progression

Strategy for Biologic Dose Reduction:

- Clinical assessment
    - DAS 28 < 2.6 or LDA +/- USS remission
    - BASDAI & Pain VAS <4 – expect to be much less than 4 and >50% improvement from pre biologic
- Reduce biologic drug by one third
- Follow up at 3 months (plus Advice Line)
- If flares – retreat at full dose
- If LDA or remission, review every 6 month, consider further reduction

Table 78: Exeter biologic clinic recommendations on dose reduction strategies

Biologic Drug	1st dose reduction	2nd dose reduction
Adalimumab	40mg every 3 weeks	40mg every 4 weeks
Etanercept	50mg every 10 days	50mg every 14 days
Certolizumab	200mg every 3 weeks	200mg every 4 weeks
Golimumab	50/100mg every 6 weeks	50/100mg every 8 weeks
Infliximab IV	2mg/kg every 8 weeks / per infusion	2mg/kg every 12 weeks / per infusion

Appendix 6. Recommendations by NHS Greater Glasgow and Clyde

Serum sample required for trough level should be taken pre-infusion for infliximab and no earlier than 3-5 days prior to injection date for adalimumab. Test results are interpreted as follows (Table 79, Figure 18 and Figure **19**):

- Levels below the lower limit suggest secondary failure of response or poor compliance. Presence of neutralising antibody may be present in the former.
- Levels above the upper limit suggest overtreatment.

Table 79: Interpretation

Analyte	Lower limit of assay	Upper limit of measurement	Units
Adalimumab	0.4	14	ug/mL
Infliximab	0.3	14	ug/mL

Figure 18: Interpretation: 3-6/12 after initiation of therapy to guide drug dose/infusion time interval

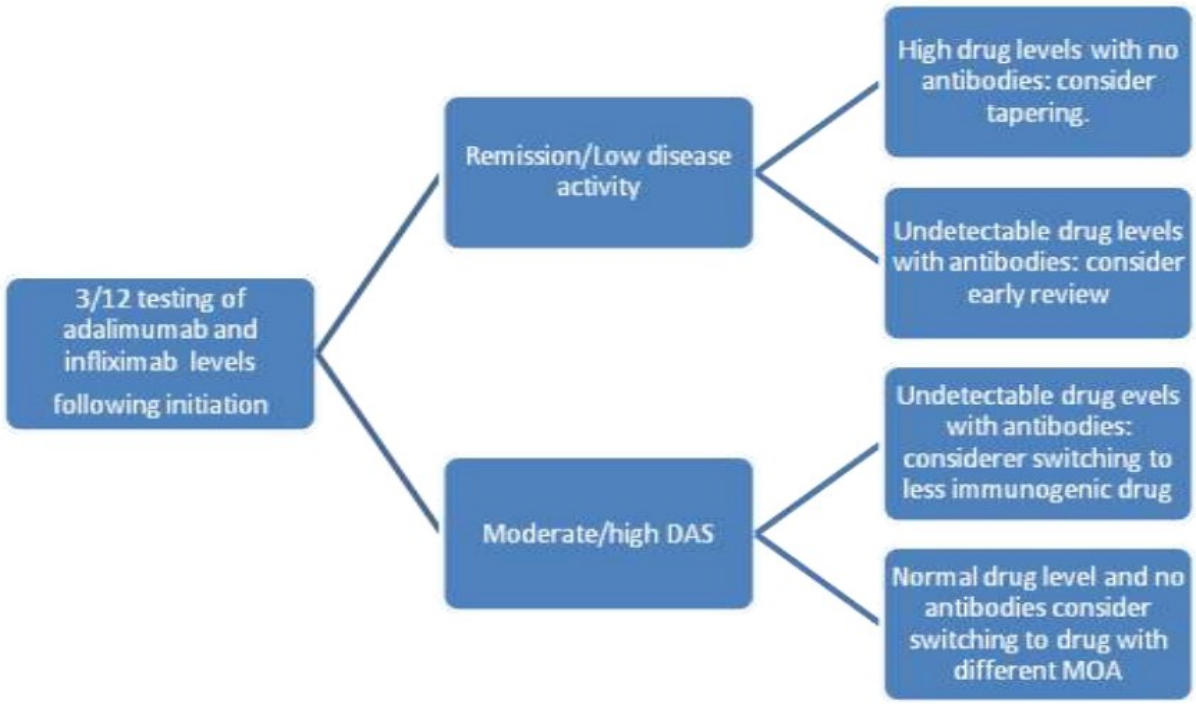
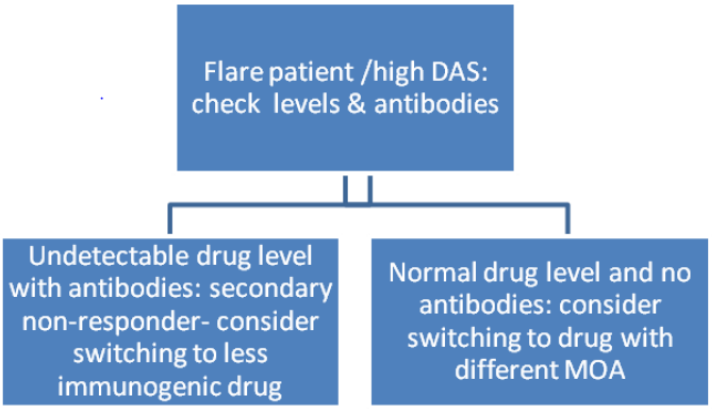


Figure 19: Interpretation: anti-TNF failure of response



Interpretation: considering dose reduction

- High/normal drug levels confer favourable likelihood of success.
- Undetectable drug levels with presence of antibodies suggest drug is not required for the patient's remission. Consider stopping therapy.

Appendix 7. NICE reference case

Table 80: Summary of the reference case

Element of health technology assessment	Reference case	Met / not met	Notes
Defining the decision problem	The scope developed by NICE	Y	
Comparator(s)	As listed in the scope developed by NICE	Y	
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Y	
Perspective on costs	NHS and PSS	Y	
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Y	
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	N	The time horizon was 18 months (due to limitations in clinical effectiveness evidence base)
Synthesis of evidence on health effects	Based on systematic review	N	Based on the only relevant study identified from the systematic review.
Measuring and valuing health effects	Health effects should be expressed in QALYs.The EQ-5D is the preferred measure of health-related quality of life in adults.	Y	
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Y	
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Y	
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Y	
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	N	All costs except the cost of managing flares were relevant to the NHS. The cost of managing flares was sources from a study conducted in France.

Element of health technology assessment	Reference case	Met / not met	Notes
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	N	Discounting was not applied because of short-term time horizon adopted in this study.

Key: NICE, National Institute for Health and Care Excellence; NHS, National Health Service; PSS, personal social services; QALYs, quality-adjusted life years; EQ-5D, standardised instrument for use as a measure of health outcome.

Appendix 8. Average cost of joint replacement surgery (the Royal Devon & Exeter NHS Foundation Trust)

Estimates related to the cost of surgery in rheumatoid arthritis patients (the Royal Devon & Exeter NHS Foundation Trust) are shown in Table 81. International classification of diseases (ICD) 10 codes included any codes from categories M05 or M06 or code M08.0 in conjunction with the OPCS procedure codes which could be any from categories W37.- W38.- W39.-W40.- W41.- W42.- W43.- W44.- W45.-W46.- W47.- W48.-W49.- W54.- W58.- W93.- W94.- W95.- W96.- W97.- W98.- O06.- O07.- O08.- O18.- O21.- O22.- O23.- O24.- O25.- O26.- O32.- Those categories are all relevant to joint replacement surgeries. The time period considered was April 2017 - September 2018. This data was provided to the AG by Nicola Finch, Leanne Brown, Keith Oldfield and Rob Storey from the Royal Devon & Exeter NHS Foundation Trust.

Table 81: Estimates related to the cost of surgery in people with rheumatoid arthritis (the Royal Devon & Exeter NHS Foundation Trust, April 2017 - September 2018)

Specialty Group Description	POD Group Description	ICD10 (Diagnostic)	OPCS (Procedure)	Cost Actual	Episode Count	Average Cost per episode
Orthopaedics	In Patients	M0596 Seropositive rheumatoid arthritis, unspecified	W401 Primary total prosthetic replacement of knee joint using cement	£7,418.75	2	£3,709.37
Orthopaedics	In Patients	M0645 Inflammatory polyarthropathy	W371 Primary total prosthetic replacement of hip joint using cement	£4,613.91	1	£6,242.08
Orthopaedics	In Patients	M0690 Rheumatoid arthritis, unspecified	O211 Primary total prosthetic replacement of elbow joint using cement	£6,351.26	1	£6,351.26
Orthopaedics	In Patients	M0691 Rheumatoid arthritis, unspecified	O071 Primary hybrid prosthetic replacement of shoulder joint using cemented glenoid component	£4,037.03	1	£4,037.03
Orthopaedics	In Patients	M0694 Rheumatoid arthritis, unspecified	W541 Primary prosthetic replacement of articulation of bone NEC	£5,388.32	1	£5,388.32
Orthopaedics	In Patients	M0696 Rheumatoid arthritis, unspecified	W401 Primary total prosthetic replacement of knee joint using cement	£33,273.88	5	£6,654.78

Specialty Group Description	POD Group Description	ICD10 (Diagnostic)	OPCS (Procedure)	Cost Actual	Episode Count	Average Cost per episode
Orthopaedics	In Patients	M0697 Rheumatoid arthritis, unspecified	O321 To be defined	£2,076.00	1	£2,076.00
Orthopaedics	In Patients	M0699 Rheumatoid arthritis, unspecified	W371 Primary total prosthetic replacement of hip joint using cement	£4,590.18	1	£4,590.18
Plastic and Reconstructive surgery	In Patients	M0694 Rheumatoid arthritis, unspecified	W541 Primary prosthetic replacement of articulation of bone NEC	£6,857.30	2	£3,428.65
Total				**£74,606.62**	**15**	**£5,061.80**

Appendix 9. Estimation of the costs of managing disease health states

Figure 20: Distributions and costs for different HAQ bands