The responses made by ScHARR-TAG to the comments made by consultees and commentators on:

”Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying anti-rheumatic drugs and after the failure of conventional disease-modifying anti-rheumatic drugs only: systematic review and economic evaluation”

October 2013

Contents Page

Preface

Factual errors identified by the manufacturers, and agreed by the Assessment Group, or identified by the Assessment Group that should be explicitly highlighted

Frequently Raised Comments

Individual Responses to consultees and commentators

References

Preface

The following criteria have been used when addressing the comments made by consultees and commentators

1) Only comments directly relating to the cost-effectiveness of the interventions will be addressed. Comments speculating on potential NICE appraisal committee decisions (both recommendations and preferred choice for parameter values) will not be addressed, comments relating to political issues or to analyses outside of the scope of this appraisal will not be addressed.
2) For readability, lengthy comments have been distilled into key points which are addressed. We have attempted to respond to all points related to the cost-effectiveness of the interventions. We have tried not to lose any key points within the abridgement.
3) Where the same comment has been raised on multiple occasions the reader will be cross-referenced to either a previous response to a specific commentator, or if the point is deemed to be extremely important, a frequently raised comment section that precedes the responses to individual consultees and commentators.
4) To reduce the length of this document, where consultees or commentators have provided additional data based on statements in the Assessment Group report, (such as correlation parameters) or have identified errors in their initial submission, these have not been discussed here unless deemed to significantly change the ICER or the Assessment Group interpretation of the ICER.
5) Similarly, points of a semantic nature which would not affect the ICERs have not been debated in this document.

This report has been divided into:

- Factual errors identified by the manufacturers and agreed by the Assessment Group that the Assessment Group deem should be explicitly highlighted
- Frequently raised comments
- Individual Responses to consultees and commentators

Factual errors identified by the manufacturers, and agreed by the Assessment Group, or identified by the Assessment Group that should be explicitly highlighted

Table 3 in the Assessment report contained typographical errors. These affected the administration cost per subcutaneous injection and the cost per year, and additional costs in year 1 for the interventions. These costs have been corrected and a revised table is provided below. Note that the next to last column now reads as cost per year (including administration costs). The administration costs were assumed to be £3.05 per subcutaneous injection, not £2.61 as stated on p354. The £3.05 cost was calculated assuming 30 minutes (including travel) of district nurse time per injection. All of the typographical errors were contained within the table and were not propagated within the cost-effectiveness analyses.

The treatment pathways for those who cannot receive MTX have been assumed to contain tocilizumab as a third-line intervention, which is outside of NICE's recommendation. This will introduce some inaccuracy, however given the similarity of the costs and efficacy of the bDMARDs the results run provide an indication of the likely ICERs for sequences (of three bDMARDs) that follow NICE guidance.

Table 3: The assumed mean acquisition costs for each intervention

Treatment	Dose regimen	Details of PAS if applicable	Cost per cheapest available dose (dose)	Cost per weight-adjusted dose [1] / standard regimen	Administration costs per treatment	Cost per Year (including admin costs) [2]	Additional Costs in Year 1
Abatacept (intravenous)	500 mg below 60 kg, 750 mg between 60-100 kg, 1000 mg above 100 kg; 0, 2 and 4 weeks then every 4 weeks thereafter	[redacted]	[redacted] (250mg)	[redacted]	£154	[redacted]	[redacted]
Abatacept (subcutaneous)	125mg weekly following loading dose 500 mg below 60 kg, 750 mg between 60-100 kg, 1000 mg above 100 kg.	[redacted]	[redacted] (125mg)	[redacted]	£3.05	[redacted]	[redacted]
Adalimumab	40 mg; every other week	N/A	£352.14 (40mg)	£352.14	£3.05	£9234.94	£-
Certolizumab pegol	400 mg per week initially, repeated at weeks 2and 4 weeks followed by a maintenance dose of 200 mg every 2 weeks	Initial 10 doses free	£357.50 (200 mg)	£357.50	£3.05	£9374.30	-£2523.85 [3]
Etanercept	50 mg; every week	N/A	£178.75 (50mg)	£178.75	£3.05	£9453.60	£-
Golimumab	50 mg below 100 kg, 100 mg above 100 kg, per month	100mg dose provided at the same price as the 50mg dose	£762.97 (50mg)	£762.97 [4]	£3.05	£9192.24	£-
Infliximab [5]	3 mg/kg: 0, 2, 6 then every 8 weeks	N/A	£419.62 (100mg)	£1110.98	£154	£8222.37 [6]	£1820.47
Tocilizumab	8 mg/kg every four weeks	[redacted]	[redacted] (80mg)	[redacted]	£154	[redacted]	£-

[1]Assuming the weight distribution of patients from the BSRBR and choosing the least expensive method of meeting the requirement. The correct dose for a specific patient is calculated within the model. [2]Assuming no vial sharing [3]This value has been simplified for clarity and is negative due to assuming 10 free doses in year 1 as detailed in the patient access scheme. The model calculates the timing and number of doses correctly. [4]Assuming that the cost of 100mg syringes are set to the price of 50mg syringes as per the previously agreed patient access scheme. [5]These values have been simplified for clarity, assuming 8 doses in year 1 and 6.5 in each subsequent year. The model calculates the timing and number of doses correctly. [6]Assuming no increase in dose requiring additional vials, - if the response is inadequate after 12 weeks, the dose may be increased in steps of 1.5 mg/kg every 8 weeks, up to max. 7.5 mg/kg every 8 weeks; alternatively, 3 mg/kg may be given every 4 weeks.

N/A – not applicable

Frequently Raised Comments

The Assessment Group Model not being released.

The Assessment Group understands the frustration of consultees and commentators that the Assessment Group model has not been released as this would have provided the exact calculations being undertaken within the estimation of the ICERs. The release of the model is not within the Assessment Group's control, and we have sent NICE a version of the model to distribute once confidentiality issues (with one of the manufacturers) have been resolved. For clarification, at present, were the model to be released, consultees and commentators could iteratively test prices for one intervention until the results matched those in the Assessment Report and would thus be able to infer a commercial in confidence discount, which the manufacturer refuses to release confidentially to consultees and commentators.

The results of the Assessment Group model are contradictory to previous estimates of the cost-effectiveness of bDMARDs.

It has been commented that the results of the Assessment Group model is discordant with previously published models and NICE technology appraisals. However, as can be seen from Table 194, when the data have been set to previous values (linear HAQ, discount rates of 6% for costs and 1.5% for QALYs) the model produces values around £23,000, albeit with a more favourable HAQ to utility mapping which do not seem greatly inconsistent with previous estimates. Thus, it appears that the criticism relates to the use of the evidence populating the model.

It is expected that a greater knowledge base on factors such as efficacy, costs and utility would cause the ICER to change over time. As seen in Table 194, the primary driver of the ICER is the assumed HAQ progression whilst on cDMARDs, for which the Assessment Group uses a more sophisticated approach and alternative datasource in the base case analysis compared with previous economic evaluations. The Assessment Group believe that the use of these data has advantages over shorter-term, smaller size RCTs which have been used previously.

Modelling should be based on ACR response rather than EULAR response

The manufacturers highlight a number of limitations with using EULAR as the response criteria rather than ACR. The Assessment Group note these limitations, but would still contend that continued treatment based on EULAR response better represents current UK practice than does an ACR response, which also has limitations.

The EULAR response criteria is reported in numerous studies, endorsed by the British and European Rheumatology Societies (BSR and EULAR), and reflects the existing NICE guidance which are based on DAS changes (the metric underpinning the EULAR response criteria). ACR is not routinely measured by clinicians in the UK. To construct a model that is based explicitly on ACR requires assumptions to be made about the relationship between ACR and clinical decisions to stop or start therapies. Since this is done using EULAR inter alia, such an approach only obscures the assumed relationship between the two outcomes. The ScHARR model is entirely explicit about this issue.

However, in order to evaluate the impact of using a EULAR, rather than ACR response separate analyses were undertaken, mapping ACR data (which is acknowledged to be more plentiful than EULAR data) onto EULAR data.

It has been commented that for consistency that ACR response should be the preferred measure. This is self-perpetuating as ACR models had typically been presented to the NICE Appraisal Committees. Observers at the STA for abatacept (for which ScHARR-TAG was the Evidence Review Group) would have heard the Appraisal Committee state a preference for a EULAR model as they believed it would better reflects current UK clinical practice.

There have been requests for the model to be rebuilt to use ACR measures directly. Due to the belief that the EULAR response criteria have far more relevance to current UK practice than ACR response this has not been done and the considerable resources required to rebuild the model this has not been undertaken.

The Assessment Group comment that the results produced by both the EULAR and ACR models are relatively similar (Table 194, p386) and that this therefore provides no indication that the typical ICERs produced by the EULAR model are discrepant with those from those using an ACR response.

HAQ progression whilst on cDMARDs

It has been commented that the data set (ERAS) used in calculating HAQ progression whilst on cDMARDs may underestimate HAQ progression as 'Only patients with highly active and rapidly disabling disease failing on cDMARDs would be considered for biological therapy, and the HAQ progression in these patients will be very different from those with milder disease who will never be considered for bDMARDs……. If the HAQ progression for cDMARDs was confined to those patients most likely to need bDMARDs, this would be more appropriate than amalgamating all RA patients." (British Society for Rheumatology) This comment does not appear to take into consideration that all four groups identified by Norton et al [1] exhibit a reduction in HAQ increase over time, regardless of the increases in the 5 year period (after initial response) that is incorporated within the Assessment Group model. This is in direct contrast to assuming that all patients will eventually have a HAQ of 3 whilst on cDMARDs only, should they live long enough as has been the approach taken by all manufacturers.

Figure 1. The classes of patients on cDMARDs estimated by Norton et al.

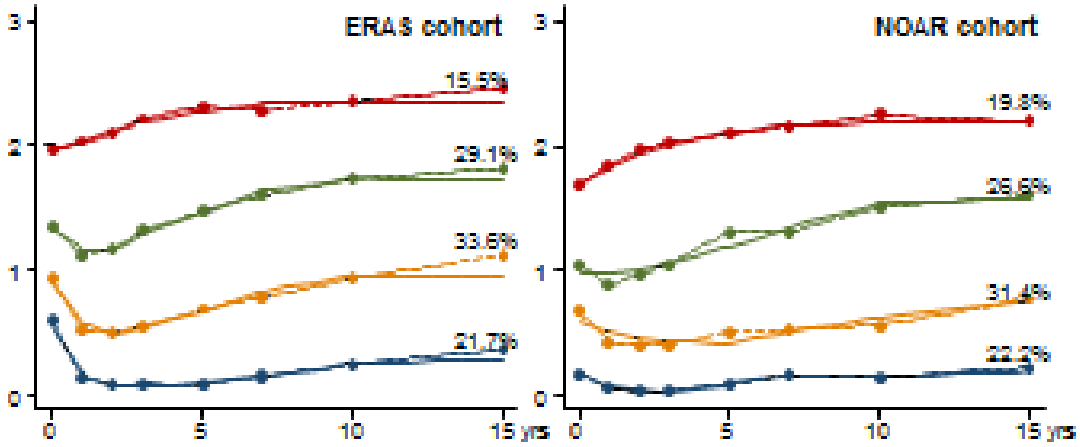

Figure 1 provides evidence of four distinct groups drawn from the entire ERAS cohort. This includes both patients that would become eligible for bDMARDs and those that would never do so, as correctly noted by the BSRBR. In order to apply the ERAS analysis to the population of interest, the analysis estimates the probability of class membership is a function of patient characteristics. Since we only sample patients with characteristics that conform to the bDMARD eligible population, the appropriate HAQ progression rate can be estimated. This process means that we sample patients much more likely to be in classes 2 and 3 (green and yellow lines). These are the classes which exhibit the worst HAQ progression in the initial ten years.

For example, the percentage of class membership for patients who are MTX-experienced, with a DAS of greater than 5.1 was: 8.77% for the red (upper) line; 40.18% for the green line; 32.00% for the yellow line and 19.05% for the blue (lower) line. This shows that patients receiving bDMARDs are more likely to be in the middle classes which exhibit the worst HAQ progression in the initial ten years, compared to the overall ERAS population. For information, the annual change in HAQ between year 2 and year 7 was 0.044 for the red line, 0.086 for the green line, 0.06 for the yellow line and 0.018 for the yellow line. Weighting these values by the predicted probability of class membership gives a value of 0.06 HAQ increase per year between years 2 and 7.

MSD undertook a systematic review of the literature (detailed in their response to the Assessment Group Report) and conclude that 'the sensitivity analysis applied by the AG of linear progression is at least as plausible as their basecase assumption'. The results of the literature review do not report: population characteristics; length of follow-up (which is crucial as the data used by the Assessment Group have high rates of decline in the initial years); and the numbers of patients within the analysis. The Assessment Group comment that the data used within their base case is UK-based, has 15 years of follow-up, included over 1300 patients with attempts made to adjust the data to represent the appropriate populations by the use of covariates. The Assessment Group therefore believe that these data are highly applicable to the decision problem.

As commented by Pfizer, the 'levelling off in HAQ seen in the analysis' could be due to 'joint surgery, exclsion arthroplasty and/ or joint fusion.' The Assessment Group agree with this statement and highlight the fact that all of the submitted models include the costs for such procedures. Failure to

incorporate the associated health benefits of surgery, as is the case by assuming a constant HAQ worsening, is inappropriate and biased in favour of bDMARDs.

The mapping of ACR response to EULAR response

The data used within the report were originally academic-in-confidence. These data can now be made public, and are contained in Table 1. These data have been assumed to represent the actual proportions and no uncertainty has been included.

Table 1: The relationship between EULAR responses and ACR responses in the VARA database

	Less	*ACR20*	*ACR50*	*ACR70*	*total*
EULAR ESR, all patients					
EULAR None	755	4	2	0	759
Mod	136	27	2	2	163
Good	57	26	10	2	83
EULAR ESR, severe active					
EULAR None	72	2	0	0	74
Mod	33	19	0	0	52
Good	3	9	5	1	12

It has been commented that the VARA database may not be generalisable to the UK as it is American data, although no reason has been put forward as to why the relationship between ACR and EULAR would change based on location. The Assessment Group deem it reasonable to assume the relationship holds, particularly if should it not, the generalisability of all RCTs conducted outside of the UK would also become debatable.

Assuming HAQ change conditional on EULAR responses but not on bDMARD vs cDMARD will be unfavourable to bDMARDs

Currently there are no data to support an assumption of either equal of differential effect in HAQ change between bDMARDs and cDMARDs conditional on EULAR response. Analyses of the relative change in HAQ conditional on ACR response presented by the manufacturers (and summarised in section 6.2.12 of the Assessment Report) shows no clear advantage for bDMARDs compared with cDMARDs.

Treatment effect on pain is not directly proportional to HAQ score and is assumed equal across all populations

All models that do not include pain assume two things, albeit implicitly:

i) That health utility can be entirely explained by HAQ score (which is equivalent to the assumption that HAQ and pain are perfectly correlated)
ii) That treatment effects therefore need only be described in terms of HAQ

We argue and provide substantial empirical evidence that i) is not the case. Better estimates of the effects of therapies on health utility must look at how HAQ and pain change.

Ideally one would conduct full evidence synthesis of to estimate the effect of therapies on pain and its correlation with HAQ. This is not feasible.

The second best solution is to use an estimate of the empirical relationship between HAQ and pain which assumes that therapies that improve HAQ have a predictable impact on pain. It does not differ due to any other patient characteristics other than to the extent to which those characteristics make patients more or less likely to be in a particular degree of functional severity. This maintains the ability to accurately predict health utility for a fully described health state. This does not double count benefits at all, but will be inaccurate if therapies have a proportionally greater or lesser impact on pain than function.

It should be noted that this risk is entirely equivalent in all other models which explicitly ignore pain in the mapping function but these are subject to eh additional bias that they are predicting utility from an inadequately description of health so have an additional layer of inaccuracy that is not present in the ScHARR approach.

But note, this would be the case if pain were to be ignored as all other models do

Use of utility and HAQ relationships from the USA may be unrepresentative of UK patients

We know of no reason why the mapping between HAQ and EQ-5D from the USA would not be generalisable to the UK. We have used the UK index scores for the EQ-5D.

The role of intensive cDMARD therapy prior to bDMARDs.

The strategies modelled by the Assessment Group in Populations 2 and 3 all assume that intensive cDMARD therapy has been provided prior to the initial bDMARD. This concurs with comments from consultees and commentators and is consistent with NICE guidance (CG79). The Assessment Group did not (for time reasons) evaluate initial intensive cDMARDs in Population 1 and employed a single cDMARD strategy followed by intensive cDMARDs. Given both the scope of the work to evaluate initial bDMARD use, the relatively high ICERs for bDMARDs in such a population and consensus from bodies such as the BSR that "most patients with active RA should try intensive DMARDs prior to starting biologics is reasonable and well argued" and "the conclusion that bDMARD therapy is highly unlikely to be cost effective in patients who are cDMARD naïve is almost certainly correct." (Healthcare Improvement Scotland) would mean that the exclusion of initial intensive cDMARDs strategies in Population 1 would not change any conclusion regarding the cost-effectiveness of bDMARDs in this group.

The exclusion of lost productivity from the report

Lost productivity has not been included in the economic evaluationas it does not constitute the NICE reference case. Secondly, were this to be robustly included, the increased lost productivity in patients who would have their treatment stopped were RA treatments to consume more resources, or the reduction in lost productivity in those patients who receive treatment, and previously would not, were the expenditure in RA interventions reduced would have to be calculated.

Unfair criticism of manufacturers who had in the Assessment Group's opinion not evaluated the correct population.

The Assessment Group believe that BMS and MSD did not provide the Appraisal Committee with information on the cost-effectiveness of Population 2 *and* of Population 3 (however defined). Where there is a possibility that two subgroups have different ICERs aggregating the two groups may result in a subgroup that is cost-effective to treat appearing to not be, or contrastingly a subgroup that is not cost-effective to treat appearing to be cost-effective to treat. Whilst the exact definitions of severe and moderate to severe populations were not provided, and could be determined by the manufacturers, it is still necessary to evaluate both populations in isolation. The cost-effectiveness of Population 3 could only be approximately inferred from data provided by MSD, and could not be disentangled from the results provided by BMS. The Assessment Group note that all but BMS and Roche (who only looked at an MTX intolerant or contraindicated population) provided analysis for Population 2 using the same definition as the Assessment Group, and that all manufacturers (excluding Roche) provided analysis for a Population with a DAS score ≥ 3.2.

Certolizumab should be associated with a greater incidence of adverse events than other bDMARDs

Data provided as academic-in-confidence by UCB (the manufacturers of certolizumab) together with wide confidence intervals for all interventions led the Assessment Group to assume that the adverse events were equal for bDMARDs. Should it be assumed these adverse events are greater for certolizumab than for other bDMARDs then the cost-effectiveness of certolizumab would become less favourable.

The network meta-analysis (NMA) has been undertaken for moderate to severe patients and severe patients combined.

The Assessment Group confirm that was done, for the reason that disaggregated data from the large number of trials which recruited patients from both populations were not available. Similar approaches were taken by some manufacturers. The Assessment Group believe that pooling the patient populations was necessary to allow a full incremental analysis to be undertaken and was not related to whether a EULAR or ACR model was being developed.

Whilst Pfizer (along with UCB) has raised this issue it is noted that in the Pfizer submission it is stated that 'A systematic review was unable to identify sufficient information to permit estimates of relative efficacy to be made within the Moderate to Severe population (Section 2). Therefore in the absence of evidence, the relative treatment effects from the Severe DMARD-IR combination therapy population were utilised. Limited evidence has observed that relative treatment effects are likely to be consistent across these populations.' Therefore Pfizer's approach does not use differential efficacy between the populations.

The inclusion of TACIT data

It has been commented that the inclusion of TACIT data could distort the results as there were 3 bDMARDs grouped together within an arm and that the RCT was of open-label design. A request was made to run sensitivity analyses without incorporating TACIT data. The Assessment Group comment that TACIT will have very little impact on the results as the trial connected Intensive DMARDs to Grouped biologics (see Figure 5 p127)and ICERs for group biologics were not evaluated. The inclusion of TACIT data does however, provide supporting evidence for the Appraisal Committee in determining the respective efficacies of intensive cDMARDs and grouped biologics, albeit with limitations regarding homogeneity of bDMARDs and an open-label RCT design.

Individual Responses to consultees and commentators

Note that some points may not be answered here, but have been answered in the frequently raised comment section.

Response to comments from Abbvie

	Distilled point	ScHARR-TAG response
1	Inclusion of unblinded studies in the NMA	The Assessment Group acknowledges the potential limitations associated with open label RCTs compared with double blind RCTs. Notwithstanding, the Assessment Group considers that the exclusion of open label RCTs would have resulted in the omission of relevant clinical trial evidence (for example, BeST, Swefot trials) and this deemed more important than any potential bias introduced. Information on blinding status was presented in Tables 343 to 347.
2	bDMARDs used as monotherapy and used in combination with a cDMARD have been assumed to be unrelated interventions	The Assessment Group has used a conventional analysis in which interventions are treated as distinct.
3	The efficacy of bDMARDs in MTX-experienced and MTX-naïve patients are assumed to be unrelated.	The Assessment Group confirm that the effects of treatments were estimated separately and dependent on previous treatment with cDMARDs
4	Populations with substantial prior exposure to bDMARDs have been included in sensitivity analyses without incorporation of meta-regression	The Assessment Group only included studies in patients with limited exposure to bDMARDS. Our belief was that this would not affect the inferences unduly and we did not investigate whether a meta-analysis would be feasible.
5	Inaccurate reporting of RACAT data.	The data used by the Assessment Group was that presented at 24 weeks (Table 2 of the NEJM report) which was the appropriate time point for our NMA.
6	Statements on relative effectiveness of bDMARDs are not justified.	The comment regarding the relative effectiveness of interventions in the two populations relates specifically to the analysis of ACR data because we did not do an analysis of EULAR data for Population 1. The probabilities of treatment rankings in the two populations will be affected by the fact that there are different treatments in the analyses. The Assessment Group did not make any claim about the relative rankings of the treatments, only which ones are most likely to be the best. We acknowledge that there is more uncertainty about which treatment is the best in Population 2/3 than Population 1.
7	Studies conducted in Asia should be excluded from the base case	The AG notes the points raised relating to the exclusively East Asian RA trials (for example limited generalisability to the UK with regards to ethnicity of trial populations and lower dose prior methotrexate

		utilised in some trials). However, discussions with our clinical advisors led to the inclusion of studies within the base case, with an intention (that was not actioned due to time restraints) to undertake sensitivity analyses without these studies. The Assessment Group comment that based on the NMA conducted by Abbvie, the inclusion of Asian trials appeared to favour tocilizumab monotherapy, and that this would be less favoured were the Asian trials removed.
8	A recently published abstract of the (open label) CONCERTO trial has been excluded from the systematic review	This abstract was published after the cut-off date for identifying literature and was therefore not included. It appears these data were first reported at the EULAR 2013 (June) conference.
9	The use of EULAR may favour tocilizumab due to a higher weighting of acute phase response than ACR	
10	The costs of infusion should be made intervention specific based on infusion duration.	In principle, infusion duration could affect infusion costs, although as a previous NICE appraisal were content to use a single value and deliberate on the potential effects the Assessment Group chose this approach rather than a potentially arbitrary division between day case and outpatient tariffs.
11	The cost of those requiring home help per subcutaneous injection (an average of £3.05 per injection) should not be applied to adalimumab as this is funded within Abbvie's Homecare provider service.	If the service is provided free of charge by Abbvie then this cost has been erroneously applied, although it is also unclear how many other subcutaneous manufacturers provide a similar service. It is not believed that the additional £3.05, on the acquisition cost of £352.14 will substantively change the results.
12	Changes in HAQ score can influence mortality and this would favour bDMARDs if incorporated (although the impact of a different assumption is not large)	The exact change in the ICER was not presented by Abbvie, but we are not surprised that the impact is not large. We believe that the assumptions used in the Assessment Group report are defensible, but agree that if it were proven that changes in HAQ did influence mortality rates then this would (slightly) improve the cost-effective of bDMARDs.
13	The costs of treating serious infections may be underestimated	It is not clear that these are under-estimated. Sensitivity analyses (p384) showed that by increasing the impacts of both costs and disutility by 100-fold have little impact on the ICERs, so it is unlikely that a minor under-estimate of costs would substantively change the results. If the costs of treating serious infections are underestimated then this will slightly disfavour bDMARDs, as would the exclusion of other adverse events.

Response to the Arthritis and Musculoskeletal Alliance

No further comment required.

Response to BMS

	Distilled point	ScHARR-TAG response
1	EULAR response is equated to a DAS28 response of ≥1.2, whereas moderate responders can improve DAS by only 0.6-1.2 if their final DAS28 is ≤5.1	This is correct, although the Assessment Group believed that the slight inaccuracy introduced in this respect was outweighed by using a response criteria much more relevant to UK clinical practice.
2	People withdrawing due to adverse events in the first 6 months should receive a second bDMARD rather than rituximab.	This is correct, although few patients fall into this category so the results are not expected to substantively change. The direction of the ICER is also unclear because rituximab has comparable efficacy to bDMARDs at half the price and thus for those patients who remain on rituximab for a considerable time this is more cost-effective than remaining on a first-line bDMARD for a considerable time.
3	Tocilizumab should not be automatically used following rituximab.	Whilst NICE recommendations do not specifically state that tocilizumab must be used post rituximab, it is recommended in this position and we believe the vast majority of clinicians would rather use tocilizumab than cDMARDs, so are not concerned that the Assessment Group assumption biases the ICERs
4	A BMS analysis attempting to implement the HAQ trajectory used by the Assessment Group increases the ICER by only 30-40% in contrast to the 100% shown in the model.	As noted by BMS their approach is simplified, particularly as they have not adjusted for covariates which predict class membership. The size of the increase is not constant due to the ICER being a ratio and the Assessment Group notes the increase is only approximately 50% when the relationship between HAQ and pain is taken from ERAS (p386).
5	Tolerability profiles are not differentiated by intervention.	BMS suggest the rate of patients discontinuing treatment be taken from RCTs. The Assessment Group has used a large database to provide markedly more observations than from the RCTs, conditional on EULAR response, and which is expected to provide a more robust estimate of discontinuation levels.
6	The EULAR responses are assumed to be equivalent for both the moderate to severe, and the severe patient groups.	We can confirm that EULAR response is assumed to be independent of whether the person is in Population 2 or 3 and apologise that this was not explicitly stated. Bar stating this limitation, BMS hypothesise that this could either favour Population 2 over Population 3 or vice versa. As such, and given the lack of data the assumption made by the Assessment Group is still felt to be reasonable.

Response to the British Society for Rheumatology (BSR) comments

	Distilled point	ScHARR-TAG response
1	People with longer disease duration may be less cost-effective to treat than those with short duration due to a greater irreversible component of HAQ. Data from the BSRBR are taken from patients with longer disease duration than those that would be receiving an initial bDMARD in clinical practice.	We confirm that the HAQ drop associated with EULAR has been taken from the BSRBR. It is unclear whether the BSR are suggesting that, if people with lesser disease duration are treated, the proportion of patients obtaining each EULAR response is likely to become more favourable to bDMARDs or whether the absolute HAQ drop per EULAR response category is suspected to be lower (or a combination of the two.) In all cases the cost-effectiveness of bDMARDs is expected to become more favourable. However, there have been no data provided on whether this would affect the rate of HAQ progression on cDMARDs. If patients with lesser disease duration also had a slower increase in HAQ on cDMARDs then this would adversely affect the cost-effectiveness of bDMARDs

Response to Healthcare Improvement Scotland

	Distilled point	ScHARR-TAG response
1	The BSRBR data was set up to measure safety and not efficacy. The timing of measures may be important in the early weeks of treatment	The main use of the BSRBR data is in estimates of treatment duration. This is unaffected by the timings of measures. Treatment response is estimated from the clinical trials, not the BSRBR so this is also unaffected by the nature of the dataset. We note that the longer term HAQ change whilst on bDMARDs is hampered by the relatively short duration of follow up HAQ measures.

Response to MSD

	Distilled point	ScHARR-TAG response
1	Failure to report keywords in Appendix 2 means that the search strategy could not be reviewed	We apologise for the omission. This is reproduced in Appendix 1 of this documents. In response to the full comment: Stage 1 strategy combined the disease with high precision RCTs and systematic reviews filters (disease AND filters) Stage 2 strategy combined the disease with the interventions and the Cochrane High Sensitivity Search Strategy (disease AND interventions AND filter)
2	Deviation from the scope for Population 1	As detailed on p51 the populations defined by NICE were mutually exclusive but not exhaustive. Our approach for resolving this was discussed with NICE and clinical experts who deemed this reasonable. MSD did not encounter this problem as Population 1 was not modelled wiithin their submissions
3	Insufficient details provided on what the Assessment Group considered appropriate imputation of the data in early escape trials	Each trial that allowed early escape was discussed by the Assessment Group and assessed for any bias in the handling of the data. For example, such bias would exist if all such patients were excluded from the reported analyses.
4	It would be useful to compare the results from multinomial and binomial results within the NMA	The Assessment Group used an ordered categorical data model rather than a binary model so that we could estimate response rates in a single coherent model.
5	The PAS for tocilizumab has been applied in the MSD model	The Assessment Group confirm that this is the case and this was wrongly stated in the Assessment Report.

Response to Pfizer

	Distilled point	ScHARR-TAG response
1	The strategies evaluated should not include the full treatment pathway.	The scope of this appraisal is purely first-line bDMARDs and thus NICE cannot make recommendations on post bDMARD interventions. As such, and with the widespread use of post-bDMARD treatments the Assessment Group believe that the full treatment pathway must be modelled, as was done by all manufacturers (including Pfizer)
2	Reductions in clinical effectiveness for subsequent lines of therapy have not been modelled.	The Assessment Group confirm that this has not been undertaken and reiterate that if this had been included this would likely be unfavourable to bDMARDs (as also discussed by Pfizer within their response)
3	The SWEFOT trial is invalid as some patients (5/128) switched to etanercept) and that 13 patients increased the frequency of infliximab from 8 to 6 weeks)	The AG acknowledges (as stated in the report of 2 year follow-up data for Swefot by van Vollenhoven et al., 2012 [Lancet vol 359, pages 1712-1720]) that 5/128 (3.9%) patients switched to etanercept for reasons of toxicity (described in Table 347 of the AG report) and 13/128 (10.1%) patients had IFX infusion frequency increased to once every 6 weeks (with increased frequency to as often as every 4 weeks permitted in the license). However, the AG would consider these proportions to be sufficiently low to permit inclusion in the NMA. Furthermore, since the data for treatment switching were drawn from the 2 year follow-up report, it is not clear to the AG at what time point these treatment adjustments were made and so it is possible that the proportion of patients switching treatment may have been even less than 3.9% and 10.1% in the NMA trial population at 23.8 weeks. In summary, the AG considers the Swefot trial to be a useful source of clinical evidence evaluating a biologic against combination cDMARDs.
4	The use of a weakly informative prior distribution in the NMA for the between study standard deviation is an indicator of poor model convergence and fit	This is not true. Firstly, convergence of the model to the target posterior distribution was achieved. Secondly, model fit cannot tell us anything about the appropriateness of the prior information. The issue was about how to deal uncertainty about the between study prior distribution with relatively few studies. It is common, as we did initially, to incorporate reference prior distributions. However, if prior distributions do not represent reasonable prior beliefs then posterior distributions will not represent reasonable posterior beliefs. When there was insufficient sample data to update the prior distribution for the between study standard

		deviation (assumed to be consistent with extreme heterogeneity) posterior distributions about treatment effects were implausibly wide. We used a weakly informative prior distribution to acknowledge reasonable heterogeneity between studies whilst expressing our belief that extreme heterogeneity is unlikely.
5	Different methods were used for HAQ progression for cDMARDs and bDMARDs	We have used different datasources to estimate these two parameters. Methods applied differ according to the characteristics of those datasourcs and the fact that we only had access to IPD for one of those datasets to conduct analyses ourselves.
6	The costs per HAQ band used in the Assessment Group model may underestimate the true costs	The reasons for the choice of values used by the Assessment Group are provided on p375 and the assessment group feel that the costs used are reasonable. If the costs chosen by Pfizer were used then this would be more favourable to bDMARDs, although the shape of the graph of costs would need to be explained as would the larger divergence from the recent results from Bansback et al.[2]
7	Time on treatment is assumed equivalent for bDMARDs and cDMARDs	The Assessment Group clarify that the survival time for cDMARDs is assumed equivalent for bDMARDs only when the EULAR response is the same. As bDMARDs would be expected to have better EULAR responses this would translate into longer survival durations, not as Pfizer interpret the text that the average survival times are identical.
8	Combining monotherapy and combination therapy in a single NMA should be avoided without controlling for concomitant use of cDMARD.	The Assessment Group were unclear on the precise point being made and could thus not answer this
9	The STAR study should be removed from the NMA as it combines combination and monotherapy and cDMARD experienced and naïve patients	The Assessment Group notes the limitations of the STAR study but deemed the relatively small percentages of no prior cDMARDs (8%) and monotherapy arms (<18%) would not overly invalidate the trial. Whether to include STAR in the main analyses or in sensitivity analyses is a subjective decision.
10	RAPID 1 and RAPID 2 should be removed from the NMA sensitivity analyses as patients with no response at weeks 12 and 14 were classed as non-responders at week 16 and MTX was not stepped up	The Assessment Group concluded that this imputation would be reasonable and that the RCTs should be included within sensitivity analyses due to the high proportion of valid data provided.
11	Tofacitinib should not be within the NMA	The Assessment Group believe that inclusion of tofacitinib increases the evidence base and these data should be included, Whether tofacitinb has marketing authorisation is not relevant. The Assessment Group comment that restrictions on the compounds considered was primarily used to reduce the time required for the literature review

12	Data from the Comet trial should be included for Population 1 (MTX naïve)	Having reviewed the data the Assessment Group confirm that the Comet trial should have been included in the NMA for Population 1
13	The PREMIER trial should be removed from the NMA as it allows for an increase in the weekly dosing of adalimumab at week 16 if an ACR response has not been observed.	The Assessment Group notes the limitations of the PREMIER study in that 29/268 (11%) of people increased dose although the timing of the increase is unknown as it was permitted any time after week 16. However, this study was included due to the large proportion of valid data contained
14	Data from LARA on the proportion of good EULAR responses have not been included in the NMA	The Assessment Group erroneously did not divide the number of responders into Good and Moderate responders.
15	No consideration of sequences where rituximab and MTX are contraindicated.	The Assessment Group has made the simplifying assumption that whether a patient can receive MTX will be known prior to the first bDMARD and this remains constant, and that all patients can receive rituximab. The Assessment Group believe that the exclusion of rituximab as an option will disfavour bDMARDs due to the lower acquisition cost of rituximab compared with typical bDMARDs.
16	Patients in the moderate to severe population may not be representative of those seen in clinical practice.	The Assessment Group acknowledges that the BSRBR database contains a relatively small number of moderate to severe patients. However, given the relative similarity of the comparability of the ICERs of moderate to severe patients with those for severe patients in the Assessment Group model and the manufacturers' models it is not expected that this causes any significant bias.
17	Are the initial patient characteristics taken from the bDMARD cohort of the BSRBR of from all patients?	Patients are sampled from the distribution of patient characteristics seen in the BSRBR for patients with a diagnosis of RA, taking a first biologic drug.
17	No reporting of model validation or external peer review	The model was reviewed by a second member of the Appraisal Team and amended as necessary. The checks taken included: extreme value analyses; auctioning the code on pen and paper; setting identical sequences and obtaining similar answers (subject to Monte-Carlo error) and general tests for direction of movement of the ICERs given changes to input parameters. The Assessment Group also notes that the ICER of approximately £23,000 when using older discount rates and a linear HAQ progression are not too dissimilar from estimates of previous models.
18	The costs of MTX has been underestimated	The Assessment Group believe that any inaccuracy in the costs of MTX would not influence the ICER given that it is assumed to be less than £50 per year, and is also used in combination with bDMARDs
19	The costs of 'palliative care' has been underestimated (and was also underestimated by Pfizer) based on suggestions from a recent Pfizer advisory board	It is possible that these costs are underestimated but no better data has been presented. It is likely thatpatients who do not receive bDMARDs will be assigned more hospitalisation costs (across time) than those patients who receive bDMARDs

20	Monitoring costs for bDMARDs should be lower than for cDMARDs	The Assessment Group used monitoring schedules reported in Malottki et al.[3] Because concomitant cDMARD (Methotrexate) is administered alongside bDMARDs (except in the bDMARD-monotherapy indications) then the cDMARD monitoring costs are included in the bDMARD treatment regimen, resulting in equal monitoring costs for bDMARDs and cDMARDs (see end of section 6.3.10 in AR). In the bDMARD-monotherapy analyses it was also assumed that patients would receive a cDMARD monitoring schedule to ensure patient safety and treatment tolerability.
21	Simulating the expected pain score associated with HAQ will not improve the relationship between HAQ and utility as in effect HAQ alone is used to estimate utility	This claim would be correct if there were a linear relationship between HAQ and pain. However, this is not the case (and is the reason why pain is a statistically significant covariate when added to any mapping function that already contains HAQ as an explanatory variable)
22	Simulating 10,000 patients may not be sufficient to distinguish fully between the different interventions.	The Assessment Group acknowledge that there remain variations within the results from two runs of 10,000 patients with deviations shown in Figure 101 (p383) with the median ICERs being £1500 divergent. However, *within this example* the efficiency frontier remained constant (being MTX, ABT iv + MTX and ETN+ MTX). This could not be calculated from the graphs as Fig 101 was erroneously labelled and had the cost per QALY data rather than cost data, This figure is corrected below. Given the large number of analyses undertaken and the time required for the literature review and NMA, 10,000 patients were deemed acceptable.

Figure 101 (corrected) **Discounted Costs from two runs of 10,000 simulated patients**

Response to Roche

	Distilled point	ScHARR-TAG response
1	Rituximab monotherapy should be modelled although this is not within its licence.	The Assessment Group does not want to model interventions outside of licence (where this has happened for Population 2 and 3 who cannot receive MTX this was in error). The Assessment Group agree with Roche that the difference in ICERs between combination and monotherapy is due to rituximab being replaced with more expensive interventions that are assumed to be of similar efficacy.
2	RCTs that contain patients with prior bDMARD exposure should not be considered within the NMA sensitivity analyses	The inclusion of such trials is subjective and hence has only been included in a sensitivity analysis. The Assessment Group considered it important not to discard evidence from RCTs where only a small proportion of patients were bDMARD experienced, however the definition of small is arbitrary. These data have been provided in order not to limit the evidence considered by the Appraisal Committee.

Response to the Royal College of Nursing

	Distilled point	ScHARR-TAG response
1	The modelling does not seem to have taken into account the costs of joint replacement	The Assessment Group confirm that the estimated costs of joint replacement are included in the model (as detailed on pages 375 and 376)

Response to UCB

	Distilled point	ScHARR-TAG response
1	RAPID 1 and RAPID 2 should be included in the base case analysis in the NMA	This issue is subjective and another manufacturer (see response to Roche above) has argued that RCTs such as RAPID 1 and RAPID2 should not be included at all. The Assessment Group has attempted to provide the Appraisal Committee with sufficient information that estimations of ICERs are available irrespective of the conclusion reached by the Appraisal Committee on the inclusion (or not) of such trials.

Appendix 1: Search Strategies

Search strategies for systematic review of clinical effectiveness

Stage 1 clinical-effectiveness searches

Stage 1 searches identified trials and systematic reviews using, where appropriate, high precision search filters. Searches were undertaken in December 2012. Medline and EMBASE searches were updated in May 2013. In addition to the searches detailed below, evidence was sought through handsearching, citation and grey literature searching and through consultation of clinical trials registers. Further details are given in Section 5.1.1 of the main report.

MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)

1948 to July 2013

Search undertaken December 2012, updated May 2013

1. exp Arthritis, Rheumatoid/

2. rheumatoid arthritis.tw.

3. 1 or 2

4. randomized controlled trial.pt.

5. randomized controlled trial.mp.

6. 4 or 5

7. 3 and 6

MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)

1948 to July 2013

Search undertaken December 2012, updated May 2013

1. exp Arthritis, Rheumatoid/

2. rheumatoid arthritis.tw.

3. 1 or 2

4. medline.tw.

5. systematic review.tw.

6. meta analysis.pt.

7. 4 or 5 or 6

8. 3 and 7

EMBASE (Ovid)

1980 to July 2013

Search undertaken December 2012, updated May 2013

1. exp rheumatoid arthritis/

2. rheumatoid arthritis.tw.

3. 1 or 2

4. double blind:.mp.

5. placebo:.tw.

6. blind:.tw.

7. 4 or 5 or 6

8. 3 and 7

EMBASE (Ovid)

1980 to July 2013

Search undertaken December 2012, updated May 2013

1. exp rheumatoid arthritis/

2. rheumatoid arthritis.tw.

3. 1 or 2

4. meta-analysis.tw.

5. systematic review.tw.

6. 4 or 5

7. 3 and 6

Cochrane Database of Systematic Reviews (Wiley Interscience)

1996 – to May 2013

Cochrane Central Register of Controlled Trials (Wiley Interscience)

1898 to May 2013

Search undertaken December 2012

"rheumatoid arthritis" OR explode Arthritis, Rheumatoid

Cumulative Index to Nursing and Allied Health Literature (EBSCO)

1982 to April 2013

Search undertaken December 2012

1. Explode Rheumatoid Arthritis (MH)
2. Rheumatoid arthritis (TX)
3. 1 or 2
4. Randomized (TX)
5. Treatment outcomes (MH)
6. Clinical trial (PT)
7. 4 or 5 or 6
8. 3 and 7

Web of Science (ISI Web of Knowledge)

1900 – present

Search undertaken December 2012

Rheumatoid arthritis (topic) AND (randomi?ed NEAR trial*) (topic)

Web of Science (ISI Web of Knowledge)

1900 – present

Search undertaken December 2012

Rheumatoid arthritis (topic) AND (systematic review* OR meta-analys?s) (topic)

Centre for Reviews and Dissemination Database of Abstracts of Reviews of Effectiveness and Health Technology Assessment (CRD DARE and HTA)

1995 – 2013

Search undertaken December 2012

Rheumatoid arthritis (all fields)

Stage 2 clinical-effectiveness searches

Stage 2 searches identified trials using high sensitivity search filters. Searches were undertaken in April 2013

MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)

1948 to July 2013

Search undertaken April 2013

1. adalimumab.af.

2. humira.af.

3. d 2e7.af.

4. d2e7.af.

5. 331731-18-1.rn.

6. etanercept.af.

7. enbrel.af.

8. 185243-69-0.rn.

9. infliximab.af.

10. remicade.af.

11. 170277-31-3.rn.

12. ta650.af.

13. ta 650.af.

14. certolizumab pegol.af.

15. cimzia.af.

16. cdp870.af.

17. 428863-50-7.rn.

18. 1132819-27-2.rn.

19. czp.af.

20. abatacept.af.

21. orencia.af.

22. 213252-14-3.af.

23. 332348-12-6.af.

24. bms188667.af.

25. bms 188667.af.

26. ctla4ig.af.

27. ctla 4ig.af.

28. golimumab.af.

29. cnto148.af.

30. cnto 148.af.

31. simponi.af.

32. 476181-74-5.af.

33. tocilizumab.af.

34. atlizumab.af.

35. actemra.af.

36. roactemra.af.

37. 375823-41-9.af.

38. tofacitinib.af.

39. xeljanz.af.

40. tasocitinib.af.

41. cp690550.af.

42. cp 690550.af.

43. 540737-29-9.af.

44. rituximab.af.

45. rituxan.af.

46. mabthera.af.

47. 174722-31-7.rn.

48. 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 or 42 or 43 or 44 or 45 or 46 or 47

49. rheumatoid arthritis.tw.

50. exp Arthritis, Rheumatoid/

51. 49 or 50

52. 48 and 51

53. randomized controlled trial.pt.

54. controlled clinical trial.pt.

55. randomized.ab.

56. placebo.ab.

57. drug therapy.fs.

58. randomly.ab.

59. trial.ab.

60. groups.ab.

61. 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60

62. exp animals/ not humans.sh.

63. 61 not 62

64. 52 and 63

MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)

1948 to July 2013

Search undertaken April 2013

1. atacicept.af.

2. 845264-92-8.rn.

3. unii-k3d9a0icq3.af.

4. uniik3d9a0icq3.af.

5. taci-fc5.af.

6. tacifc5.af.

7. taci-ig.af.

8. taciig.af.

9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8

10. rheumatoid arthritis.tw.

11. exp Arthritis, Rheumatoid/

12. 10 or 11

13. 9 and 12

14. randomized controlled trial.pt.

15. controlled clinical trial.pt.

16. randomized.ab.

17. placebo.ab.

18. drug therapy.fs.

19. randomly.ab.

20. trial.ab.

21. groups.ab.

22. 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21

23. exp animals/ not humans.sh.

24. 22 not 23

25. 13 and 24

EMBASE (Ovid)

1980 to July 2013

Search undertaken April 2013

1. adalimumab.af.

2. humira.af.

3. d 2e7.af.

4. d2e7.af.

5. 331731-18-1.af.

6. etanercept.af.

7. enbrel.af.

8. 185243-69-0.af.

9. infliximab.af.

10. remicade.af.

11. 170277-31-3.af.

12. ta650.af.

13. ta 650.af.

14. certolizumab pegol.af.

15. cimzia.af.

16. cdp870.af.

17. cdp 870.af.

18. 428863-50-7.af.

19. 1132819-27-2.af.

20. czp.af.

21. abatacept.af.

22. orencia.af.

23. 213252-14-3.af.

24. 332348-12-6.af.

25. bms188667.af.

26. bms 188667.af.

27. ctla4ig.af.

28. ctla 4ig.af.

29. golimumab.af.

30. cnto148.af.

31. cnto 148.af.

32. simponi.af.

33. 476181-74-5.af.

34. tocilizumab.af.

35. atlizumab.af.

36. actemra.af.

37. roactemra.af.

38. 375823-41-9.af.

39. tofacitinib.af.

40. xeljanz.af.

41. tasocitinib.af.

42. cp690550.af.

43. cp 690550.af.

44. 540737-29-9.af.

45. rituximab.af.

46. rituxan.af.

47. mabthera.af.

48. 174722-31-7.af.

49. 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 or 42 or 43 or 44 or 45 or 46 or 47 or 48

50. exp rheumatoid arthritis/

51. rheumatoid arthritis.tw.

52. 50 or 51

53. 49 and 52

54. random$.tw.

55. clinical trial$.mp.

56. exp health care quality/

57. 54 or 55 or 56

58. 53 and 57

EMBASE (Ovid)

1980 to July 2013

Search undertaken April 2013

1. atacicept.af.

2. 845264 92 8.af.

3. unii-k3d9a0icq3.af.

4. uniik3d9a0icq3.af.

5. taci-fc5.af.

6. tacifc5.af.

7. taci-ig.af.

8. taciig.af.

9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8

10. exp rheumatoid arthritis/

11. rheumatoid arthritis.tw.

12. 10 or 11

13. 9 and 12

14. random$.tw.

15. clinical trial$.mp.

16. exp health care quality/

17. 14 or 15 or 16

18. 13 and 17

Stage 3 clinical-effectiveness searches

Stage 3 searches identified studies of adverse events. Searches were undertaken in July 2013. In addition to the searches detailed below evidence was sought through consultation of the National Library of Medicine (NLM) TOXLINE resource and through the website of the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA).

MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)

1948 to July 2013

Search undertaken July 2013

1. adalimumab.af.

2. humira.af.

3. d 2e7.af.

4. d2e7.af.

5. 331731-18-1.rn.

6. etanercept.af.

7. enbrel.af.

8. 185243-69-0.rn.

9. infliximab.af.

10. remicade.af.

11. 170277-31-3.rn.

12. ta650.af.

13. ta 650.af.

14. certolizumab pegol.af.

15. cimzia.af.

16. cdp870.af.

17. 428863-50-7.rn.

18. 1132819-27-2.rn.

19. czp.af.

20. abatacept.af.

21. orencia.af.

22. 213252-14-3.af.

23. 332348-12-6.af.

24. bms188667.af.

25. bms 188667.af.

26. ctla4ig.af.

27. ctla 4ig.af.

28. golimumab.af.

29. cnto148.af.

30. cnto 148.af.

31. simponi.af.

32. 476181-74-5.af.

33. tocilizumab.af.

34. atlizumab.af.

35. actemra.af.

36. roactemra.af.

37. 375823-41-9.af.

38. tofacitinib.af.

39. xeljanz.af.

40. tasocitinib.af.

41. cp690550.af.

42. cp 690550.af.

43. 540737-29-9.af.

44. rituximab.af.

45. rituxan.af.

46. mabthera.af.

47. 174722-31-7.rn.

48. atacicept.af.

49. 845264-92-8.rn.

50. unii-k3d9a0icq3.af.

51. uniik3d9a0icq3.af.

52. taci-fc5.af.

53. tacifc5.af.

54. taci-ig.af.

55. taciig.af.

56. 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 or 42 or 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

57. rheumatoid arthritis.tw.

58. exp Arthritis, Rheumatoid/

59. 57 or 58

60. 56 and 59

61. (ae or co or de).fs.

62. (safe or safety or side effect* or undesirable effect* or treatment emergent or tolerability or toxicity or adrs or (adverse adj2 (effect or effects or reaction or reactions or event or events or outcome or outcomes))).tw.

63. 61 or 62

64. 60 and 63

EMBASE (Ovid)

1980 to July 2013

Search undertaken July 2013

1. adalimumab.af.

2. humira.af.

3. d 2e7.af.

4. d2e7.af.

5. 331731-18-1.af.

6. etanercept.af.

7. enbrel.af.

8. 185243-69-0.af.

9. infliximab.af.

10. remicade.af.

11. 170277-31-3.af.

12. ta650.af.

13. ta 650.af.

14. certolizumab pegol.af.

15. cimzia.af.

16. cdp870.af.

17. cdp 870.af.

18. 428863-50-7.af.

19. 1132819-27-2.af.

20. czp.af.

21. abatacept.af.

22. orencia.af.

23. 213252-14-3.af.

24. 332348-12-6.af.

25. bms188667.af.

26. bms 188667.af.

27. ctla4ig.af.

28. ctla 4ig.af.

29. golimumab.af.

30. cnto148.af.

31. cnto 148.af.

32. simponi.af.

33. 476181-74-5.af.

34. tocilizumab.af.

35. atlizumab.af.

36. actemra.af.

37. roactemra.af.

38. 375823-41-9.af.

39. tofacitinib.af.

40. xeljanz.af.

41. tasocitinib.af.

42. cp690550.af.

43. cp 690550.af.

44. 540737-29-9.af.

45. rituximab.af.

46. rituxan.af.

47. mabthera.af.

48. 174722-31-7.af.

49. atacicept.af.

50. 845264 92 8.af.

51. unii-k3d9a0icq3.af.

52. uniik3d9a0icq3.af.

53. taci-fc5.af.

54. tacifc5.af.

55. taci-ig.af.

56. taciig.af.

57. 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 or 42 or 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

58. exp rheumatoid arthritis/

59. rheumatoid arthritis.tw.

60. 58 or 59

61. 57 and 60

62. (safe or safety or side effect* or undesirable effect* or treatment emergent or tolerability or toxicity or adrs or (adverse adj2 (effect or effects or reaction or reactions or event or events or outcome or outcomes))).tw.

63. 61 and 62

Search strategies for systematic review of cost effectiveness

In addition to the searches detailed below, evidence was sought through handsearching and grey literature searching. Further details are in Section 6.1.

MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (Ovid)

1948 to July 2013

Search undertaken January 2013

1. exp Arthritis, Rheumatoid/

2. rheumatoid arthritis.tw.

3. 1 or 2

4. Economics/

5. exp "Costs and Cost Analysis"/

6. Economics, Dental/

7. exp Economics, Hospital/

8. Economics, Medical/

9. Economics, Nursing/

10. Economics, Pharmaceutical/

11. (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).tw.

12. (expenditure$ not energy).tw.

13. value for money.tw.

14. budget.tw.

15. or/4-14

16. ((energy or oxygen) adj cost).tw.

17. (metabolic adj cost).ti,ab.

18. ((energy or oxygen) adj expenditure).tw.

19. or/16-18

20. 15 not 19

21. letter.pt.

22. editorial.pt.

23. historical article.pt.

24. or/21-23

25. 20 not 24

26. Animals/

27. Humans/

28. 23 not (23 and 24)

29. 25 not 28

30. "Value of Life"/ec [Economics]

31. quality-adjusted life years/

32. exp models, economic/

33. 30 or 31 or 32

34. 29 or 33

35. 3 and 34

EMBASE (Ovid)

1980 to July 2013

Search undertaken January 2013

1. exp rheumatoid arthritis/

2. rheumatoid arthritis.tw.

3. 1 or 2

4. health economics/

5. exp economic evaluation/

6. exp "health care cost"/

7. exp pharmacoeconomics/

8. 4 or 5 or 6 or 7

9. (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).tw.

10. (expenditure$ not energy).tw.

11. (value adj2 money).tw.

12. budget$.tw.

13. 9 or 10 or 11 or 12

14. 8 or 13

15. letter.pt.

16. editorial.pt.

17. note.pt.

18. 15 or 16 or 17

19. 14 not 18

20. (metabolic adj cost).tw.

21. ((energy or oxygen) adj cost).tw.

22. ((energy or oxygen) adj expenditure).tw.

23. 20 or 21 or 22

24. 19 not 23

25. exp animal/

26. exp animal experiment/

27. nonhuman/

28. (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep).tw,sh.

29. 25 or 26 or 27 or 28

30. exp human/

31. exp human experiment/

32. 30 or 31

33. 29 and 32

34. 29 not 33

35. 24 not 34

36. 3 and 35

Centre for Reviews and Dissemination NHS Economic Evaluation Database and Health Technology Assessment (CRD NHS EED and HTA)

1995 – 2013

Search undertaken January 2013

Rheumatoid arthritis (all fields)

EconLIT (Ovid)

1961 – 2013

Search undertaken January 2013

Rheumatoid arthritis.tw.

Web of Science (ISI Web of Knowledge)

1900 – present

Search undertaken January 2013

Rheumatoid arthritis (topic) AND (economic OR cost OR costs) (topic)

References

[1] Norton, S., Verstappen, S.M.M., Symmons, D.P., Lunt, M., Davies, R., Scott, D.L., Deighton, C., Wailoo, A.J., Tosh, J.C., Young, A. Common Trajectories Of Haq Disability Progression Over 15-Years In The Early Rheumatoid Arthritis Study And The Norfolk Arthritis Register. 2012; available from http://www.abstracts2view.com/eular/view.php?nu=EULAR12L_SAT0113

[2] Impact of Biologics on Healthcare Utilization in Patients with Rheumatoid Arthritis: An Instrumental Variable Approach Bansback N, Fu E, Sun H, Guh D, Zhang W, Lacaille D, Anis AH. Poster CRA 2013

[3] Malottki, K., Barton, P., Tsourapas, A., Uthman, A.O., Liu, Z., Routh, K. et al. Adalimumab, etanercept, infiximab, 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)