AbbVie's Comments on the Health Technology Assessment Report Prepared by the School of Health and Related Research, University of Sheffield

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 (review of TA guidance 130, 186, 224, 234, and part review of TA guidance 225 and 247)

Detailed Comments on the Assessment Report

The following sections provide a more detailed analysis of the assessment report. We have not been able to comment extensively on the economic model due to the lack of access to the model itself and uncertainty over how key inputs have been used. AbbVie considers that this limits the individual consultees' ability to comment on the assessment group modelling and compare it to the modelling of the other manufacturers.

1. Issues with the Systematic Clinical Review, Mixed Treatment Comparison and Clinical Effectiveness

1.1 The design of the Mixed Treatment Comparison using EULAR response may result in biases in the estimates of effectiveness

1.1.1 Model should be based on ACR response rather than EULAR response

The Assessment Group (AG) opted to assess the response to treatment based on EULAR response levels, and obtained estimates for response based on a Mixed Treatment Comparison (MTC) of studies which collected EULAR responses. AbbVie has the view that this option for designing the MTC results is a source of substantial uncertainty in the estimates resulting from the MTC, and that this would be alleviated if a model based on ACR responses had been adopted. Additional to this, issues relating to inclusion of studies and the incorrect consideration of control variables in the MTC estimates may promote biases and uncertainties in the estimates arising from the MTC. Finally, it is not entirely clear if the trials data in the evidence tables produced by AG has been entirely accurately reported.

The assessment of response to treatment in AG's model is based upon EULAR response at six months. The AG report states (p.345) that:

"None of the models submitted by the manufacturers replicated the clinical reality within England and Wales to the satisfaction of the Assessment Group. Primarily this is because the majority of models assumed that the efficacy of the intervention was based on improvements in ACR, whereas NICE guidance has defined stopping rules where an intervention is stopped unless a DAS28 reduction of 1.2 points is achieved. The criterion of achieving a 1.2 point reduction in DAS is associated with a good or moderate EULAR response."

1.1.1.1 DAS28-ESR and DAS28-CRP should not be used interchangeably in the assessment of response to treatment

AbbVie points out that performing a EULAR based rather than ACR based MTC poses some fundamental problems due to the fact that DAS28-ESR and DAS28CRP should not be used interchangeably in the assessment of response to treatment.

The AG report states on p. 64 that:

'It was assumed that DAS28-CRP and DAS28-ESR are interchangeable where only one is reported. If both were reported, DAS28-ESR was used as this was reported most regularly (a decision made in consultation with clinical advisors)'

There are four validated formulae in use to calculate the DAS28 score; two using the ESR and two using the CRP. The difference stems from the number of variables used in each of the formulae (three or four):

- DAS28-ESR(3)

- DAS28-ESR(4)
- DAS28-CRP(3)
- DAS28-CRP(4)

The majority of pivotal RA trials involving biologics were conducted using the four variable DAS28 scores as endpoints; however, there is inconsistency in the use of DAS28-ESR or DAS28-CRP between them. The numerical definitions of low, moderate, and high disease activity are conventionally the same regardless of which DAS28 formula is used. However, many studies examining differences between the scores using CRP and ESR have disagreed as to whether values obtained are indeed interchangeable. While some studies have demonstrated near-equivalence of the two scores[1,2,3] others have concluded that DAS28-ESR values may be significantly different and usually higher, in distinct populations; namely females, African Americans, and patients with chronic RA[4,5,6,7,8]. Lower numerical definitions for low/moderate/high disease activity have therefore been suggested for the DAS28-CRP score[5,6,7,9] although the proposed conversions between the two composite disease measures have varied significantly[5,6].

In an analysis of DAS28-CRP vs. DAS28-ESR values using data from abatacept Randomized Controlled Trials (RCTs), approximately 20% of patients (154 out of 752) were classified in a lower DAS28 category using the DAS28-CRP (e.g. moderate disease activity vs. high disease activity), of whom 49% were classified in a better EULAR response category (e.g. low vs. moderate EULAR response)[9]. A recent study[10] found a difference of 0.8 units between mean DAS28-ESR and DAS28-CRP values in African Americans (4.7 vs. 3.9 respectively; $p < 0.001$), with EULAR response based on DAS28-CRP resulting in a lower disease activity category in 40% of patients compared with EULAR response based on the DAS28-ESR.

While the interchangeable use of CRP and ESR is also a feature of the ACR response criteria; the acute phase reactant makes the lowest contribution to the total ACR score which is comprised of five core data set variables (namely: physician global assessment, patient global assessment, Health Assessment Questionnaire (HAQ), pain VAS, and acute phase reactant)[11]. This is in contrast to the high weighting of the ESR or CRP values within the DAS28 scores[6], upon which the EULAR response criteria draw to a large extent. This implies that the responsiveness of DAS28 to the choice of measurements based on ESR or CRP may place an individual in a different EULAR response category more often than it may affect an ACR response classification. It would be useful to assess this issue using the same trial dataset to observe whether this difference in classification occurs more frequently when using EULAR response criteria in favour of ACR response criteria.

1.1.1.2 Reporting of responses in clinical trials; sparseness of evidence network based on EULAR response

ACR responses have been much more frequently collected as primary endpoints in available RCTs than EULAR responses have. The EULAR responses by contrast are often not measured in RCTs or measured as secondary endpoints only. Therefore; ACR responses will have been presented more often in the literature than EULAR responses, making an MTC based on EULAR responses comparatively more prone to bias in the resulting estimates due to initial reporting bias. It should also be pointed out that more recent trials will have more likely included EULAR as endpoint, whereas ACR was collected for both more and less recent trials, which creates a time induced bias.

Furthermore the AG Mixed Treatment Comparison uses evidence networks that are quite sparse (i.e. very few comparisons between treatments are directly informed by more than one trial). In the general context of indirect comparisons, it is recognized that analyses derived from sparse evidence networks are "profoundly underpowered" and can produce results that are unreliable, particularly when there is substantial treatment effect heterogeneity between studies[12]. Indeed, the evidence network was so sparse in the case of the EULAR analyses that the AG was forced to use an informative prior distribution for the between-study standard deviation because "there were relatively few studies and too few to update the between study standard deviation" (Draft Assessment Report, p. 67). The same study also reports substantial risk of overestimation of synthesized treatment effects when only one trial informs each comparison. This risk declines as the number of studies per comparison increases. To the extent that comparisons involving adalimumab (ADA) are more densely connected to the evidence network than other treatments, it is possible that the comparative effectiveness of ADA is less likely to be overestimated than other treatments, resulting in a potential bias when comparing ADA to those other treatments. The adoption of this approach may have resulted in estimates that are less robust than a more comprehensive evidence synthesis approach. Recent research by Soares and colleagues[13] suggests that when faced with a sparse network, analysts should consider seeking model specifications and other approaches that allow for the broadest evidence synthesis possible.

1.1.1.3 Disconnect between response measures in the model and clinical practice in the UK

Finally, while it is true that achieving a 1.2 reduction in DAS is associated with a good or moderate EULAR response, and is (p. 43 of the AG report)

"more closely aligned to the treatment continuation rules stipulated by NICE that require a DAS28 improvement of more than 1.2 to continue treatment"

AbbVie points out that, whilst being based on EULAR responses, the AG's MTC does not distinguish between the ways a moderate response can be achieved. An improvement in the DAS28 score of 1.2 or more can move a patient from being above 5.1 to below 5.1; however it can also move a patient from the 3.2-5.1 category into the <3.2 category. Thus identifying moderate EULAR responders in this way can cause problems when assigning utility improvements to patients with a moderate response. Essentially this means that the same utility improvement could be assigned to patients in two very different disease categories, something which seems counterintuitive. These two types of moderate response should not be combined when conducting an MTC.

This is not a problem with the ACR20 response criterion since to qualify for such a response; all patients must exhibit a 20% improvement, regardless of where they are on the disease spectrum as a result.

1.1.2 Inclusion of un-blinded studies in the MTC

AbbVie also wishes to point out the fact that the MTC syntheses spanned RCTs and non-blinded studies and that this is an additional source of uncertainty. Differences in study design are a source of heterogeneity leading to a further source of uncertainty in the resulting estimates.

One such instance has been the inclusion of data from a phase IIb trial, IIbCREATE[14,] which boasts an un-blinded etanercept arm, and this has been used alongside data for etanercept taken from blinded studies. This could potentially introduce bias into the overall results.

A second instance, again relating to a study boasting un-blinded etanercept treatment, was also included in such a manner. It is also noteworthy to point out that this study, LARA[15], has been included as a conference abstract but referenced poorly and as such the source remains unidentifiable to an outside party. The results of the etanercept + MTX arm vs. the DMARD + MTX arm presented in this study also indicate statistical significance in favour of the etanercept arm. Again we would question the inclusion of data derived from un-blinded treatment arms alongside those that have been blinded in a base case analysis.

Finally, the AUGUST II trial[16] also contains an open ADA arm, which again has the potential to introduce bias into the estimates produced in the MTC.

AbbVie would urge caution on the inclusion of un-blinded data in this manner, and would suggest the inclusion of these as a sensitivity analysis would be more appropriate.

1.1.3 Model specification issues in the AG's MTC

Apart from the choice of a EULAR rather than an ACR evidence synthesis, and the inclusion of non-blinded studies in the evidence synthesis, several other issues should be put forward which highlight strengths of AbbVie's approach with respect to the methodology applied by the AG.

1.1.3.1 Specification issues affecting treatment effects in the AG's MTC model

First, rather than treating the use of concomitant MTX as a treatment effect modifier (as in the work of Nixon and colleagues[17]), the AG approach considers any given bDMARD as monotherapy and the same bDMARD as combination therapy to be two entirely different treatments. Therefore, no information is borrowed between mono- and combination therapy in the assessment of bDMARD performance relative to cDMARDs. Our random coefficient meta-regression analysis strongly suggests an additive (on the log-odds scale) effect of MTX for bDMARDs in general, with some heterogeneity (specifically, the additive effect of MTX seems to be larger for ADA than for other bDMARDs). The RA MTC by Nixon and colleagues [17] also argues for the treatment of MTX as a treatment effect modifier rather than as delineating separate treatments.

Second, the AG approach conducts entirely separate analyses in MTX-experienced and MTX-naïve populations (with unclear treatment of trials in mixed populations). This suggests that relative performance of DMARDs in an MTX-experienced population offers no information whatsoever about relative performance in an MTX-naïve population (and vice versa). Results obtained using our approach suggests a relatively small effect of prior MTX exposure on treatment effect, once one also accounts for concomitant use of MTX in combination therapy. Given the paucity of trials in the MTX-naïve population (and the relatively higher proportion of monotherapy trials), this failure to borrow information across MTX-experienced and MTX-naïve studies is potentially consequential.

Third, the AG approach excludes trials involving populations having substantial prior exposure to bDMARDs (but then includes such trials in sensitivity analysis, without accounting for their inclusion through meta-regression). Our

results suggest inclusion of patients with prior bDMARD exposure results in a relatively consistent reduction in bDMARD effectiveness (with a somewhat smaller effect for RTX), suggesting that including trials from prior bDMARD exposed populations can contribute information but should be controlled for.

1.1.3.2 Lack of transparency in the mapping of ACR response to EULAR response

Finally, it should be noted that the method for mapping from ACR to EULAR used by the AG is not transparent and therefore AbbVie is unable to understand what effect this has on the estimates of effectiveness of the different therapy options and resulting estimates of cost effectiveness obtained in the AG model (see Section 3.iii).

1.1.4 Inaccurate reporting of data used in MTC

AbbVie notes that there seemed to be some inconsistencies between the RACAT[18] data presented in Table 20 (p. 117-8 of the report) and the data available in the online supplement to the trial publication, namely the "n" values given for the ACR20 and ACR50 responses in the ETN+ arm; see Table 1 below.

Table 1: Inconsistency between published RACAT data and that used by the AG

	Data from p.11 of supplement		Calculated n values		Adjusted n values*		Data from AG Table 20, p.118	
	Triple DMARD	Etn+MTX	Triple DMARD	Etn+MTX	Triple DMARD	Etn+MTX	Int cDMARD	ETN+MTX
N	153	154	153	154	153	154	159	163
ACR20	58%	56%	89	86	48	31	48	32
ACR50	27%	36%	41	55	33	29	33	32
ACR70	5%	17%	8	26	8	26	8	26
No Response**	42%	44%	64	68	64	68	70	73

**Allowing for mutually exclusive response categories as per AG methodology, **Calculated as (1-ACR20) response*

Clarification on how these numbers were reached by the AG is requested. If this difference can't be justified and is in fact an error on the part of the AG; then this raises the question of data integrity throughout the tables.

1.2 Role of Combination DMARD Therapy

AbbVie points out that some of the sequences modelled by the AG are not in line with the NICE clinical guidelines and rule out improvements in effectiveness of biologics that would have been achieved if the correct sequences were modelled.

We believe that optimal treatment of RA should be in line with NICE guidance[19], including early combination DMARD therapy for all new RA patients in preference to step-up therapy from methotrexate (MTX) to combination DMARDs. However, we believe after failure of initial combination therapy, sequential switching to other DMARD combinations or additional step up should not delay commencement of biologics in combination with MTX. This is because there are a number of studies showing differences in outcomes in favour of biologics in patients with inadequate response to MTX who require escalation of therapy.

The SWEFOT trial [20] in patients with early RA and an inadequate response to MTX assessed step-up therapy to triple DMARD (MTX+SSZ+HCQ) or infliximab (IFX) combined with MTX. A significant difference favourable to IFX was found in the proportion of patients achieving a good

EULAR response at 6, 9, and 12 months, suggesting that for patients who do not respond sufficiently to MTX monotherapy, addition of an anti-TNF agent is superior to addition of additional conventional DMARDs (cDMARDs). By two years, although there was no significant statistical difference in the proportion of patients achieving a good EULAR response (38% vs. 31%; p=0.204), radiological progression was significantly greater in patients receiving triple DMARDs than biologic combination therapy (mean 7.23 vs. 4.00; p=0,009), supporting step up to biologic combination therapy in preference to combination DMARD therapy after inadequate response to MTX. It is important to remember that two years is a limited time for RA which is a chronic disease requiring lifetime treatment. Differences in radiographic progression can lead to differences in long term disability over 10 or more years.

The recent TEAR trial[21] compared 4 different treatment strategies in patients with early RA:

- 1. immediate treatment with etanercept (ETA) +MTX
- 2. immediate triple DMARDs
- 3. step-up MTX → ETA+ MTX at week 24
- 4 .step-up MTX → triple DMARDs at week 24

Although caution must be exercised with interpretation of this study due to the statistical methodology (the analysis was performed for observed data without imputation for missing values), the TEAR trial showed no difference in the mean DAS28-ESR scores between weeks 48-102 in subjects randomized to receive MTX plus ETA and those randomized to triple therapy, regardless of whether they received immediate combination or step-up from MTX monotherapy. However, a significantly higher proportion of patients in the two biologics arms compared with the triple therapy arms achieved a more robust outcome measure of ACR 70 response by year 2 (18.2% versus 11.3%; *P* = 0.01). In addition, the two groups receiving ETA had a smaller increase from baseline in Sharp scores compared with those receiving triple therapy (0.64 versus 1.69; P=0.047). Again, this demonstrates benefit of progressing to biologic combination rather than triple DMARD therapy after MTX failure. This is particularly relevant as results from the study by van Vollenhoven and colleagues 22 analysing data PREMIER trial (a 2-year, randomized controlled trial of methotrexate-naive patients with early RA (<3 years) who received treatment with adalimumab plus methotrexate, adalimumab, or methotrexate), who collected and analysed self-reported patient reports on absenteeism (number of days missed or unfit to work), presenteeism and employment status provided evidence that over 2 years, patients who received combination therapy had improved outcomes with respect to patients who received methotrexate monotherapy. They missed approximately half as many days as patients who received methotrexate (17.4 versus 36.9 days for employed workers; 7.9 versus 18.6 days for homemakers), had lower presenteeism (reflecting better productivity), and had greater likelihood of gaining/retaining employment (odds ratio 1.530, 95% confidence interval 1.038-2.255; P = 0.0318).

To conclude AbbVie considers that the treatment paradigm recommended in the NICE clinical guidelines of early first line combination therapy with conventional DMARDs is likely to be the most cost effective first line treatment. However, once patients have failed combination DMARD therapy they should be considered for anti-TNF therapy as the best way of avoiding long term disability. AbbVie is not aware of any data that shows a good response for trying a different combination DMARD therapy after failure of initial combination DMARD therapy.

1.3 Effectiveness estimates of biologic treatments

The evidence of differential relative effectiveness of Infliximab across the group of biologics between populations is not supported by real world evidence.

AbbVie questions the validity of the argument that in population 1 the use of infliximab + MTX was deemed likely to be the most effective intervention whereas in populations 2 and 3 the

treatment most likely to be effective was identified as etanercept + MTX, with infliximab + MTX ranking further down the list of efficacy. The fact that estimates of ACR response coming from the AG's MTC are more favourable to ETN used in monotherapy with respect to ADA used in monotherapy are also put into question.

These discrepancies are more likely to be attributable to differences in the populations in the RCTs included in the analysis rather than differences in the relative effectiveness of the different therapies. Such patterns are not supported by real world data.

While we are unaware of any real life data comparing the anti-TNF agents in MTX-naïve patients (due to national restrictions on use of biologics in this population), registry data in patients with incomplete response to DMARD therapy have shown IFX to have equivalent or inferior effectiveness to adalimumab and etanercept and no superior effectiveness of etanercept over adalimumab.

The DANBIO [23] Danish RA registry published outcomes from 2,326 RA patients commencing first biologic (29% received adalimumab, 22% received etanercept, and 49% received infliximab). The odds ratios (ORs) for clinical responses were calculated, corrected for age, disease duration, DAS28, seropositivity, concomitant methotrexate and prednisolone, number of previous disease-modifying drugs, centre, and HAQ score.

LUNDEX-corrected effectiveness responses at 6 months were as follows:

- An ACR70 response was achieved by 19% receiving adalimumab, 17% etanercept, and 11% infliximab
- EULAR good responses were achieved by 41%, 34%, and 27% respectively.
- DAS28 remission was achieved by 26%, 21%, and 17% respectively
- 4. CDAI remission was achieved by 15%, 10%, and 8% respectively

After correction for the differences listed above, the ORs for achieving an ACR70 response after 6 months of treatment were 2.05 for adalimumab and 1.78 for etanercept, with infliximab as the reference drug. There was no significant difference between adalimumab and etanercept (OR 1.15, with etanercept as the reference).

The ORs for achieving other treatment responses (including EULAR good response, DAS28 remission, and CDAI remission) for adalimumab versus infliximab ranged from 1.78 to 2.76 and were statistically significant for all outcome measures. For etanercept versus infliximab, the ORs ranged from 1.16 to 1.99 and were statistically significant for all outcomes except DAS28 remission and CDAI remission. For adalimumab compared with etanercept, the ORs ranged from 1.15 to 1.58 and were significant for a EULAR good response and CDAI remission. The ORs at 12 months were similar.

The DREAM registry [24] has evaluated the effects of adalimumab, etanercept and infliximab on disease activity, functional ability and quality of life and the medication costs within the Dutch Rheumatoid Arthritis Monitoring (DREAM) register starting on a first anti-TNF agents were monitored and assessed every 3 months. The primary outcome was the DAS28 course over the 12 months, analysed by linear mixed models. Secondary outcomes were the HAQ, EQ-5D, and SF36 scores, and medication-related total costs.

The repeated measures analyses (linear mixed model) showed significant differences in the DAS28 course over 12 months between infliximab and both adalimumab and etanercept patients ($p<0.001$), in favour of etanercept and adalimumab. There was also a significant difference in the DAS28 course over 12 months between adalimumab and etanercept patients ($p=0.031$), in favour of adalimumab. The figures for the HAQ and the EQ-5D showed the same

trend: the functionality and quality of life was best for adalimumab patients. The repeated measures analyses did not show any significant differences.

Pragmatic studies have been performed to compare outcomes between anti-TNF agents in the absence of randomised, double blind, head to head trials. The Red Sea trial [25] has compared effectiveness of etanercept and adalimumab. The trial was a 52 week un-blinded randomised non-inferiority trial of adalimumab versus etanercept in 120 patients with active RA despite prior or current use of two DMARDs including methotrexate (unless contraindicated). Effectiveness measures were assessed as secondary outcomes, and included DAS28 scores and health utility determined by the EQ-5D. There were no constraints on changes in the dose of methotrexate, use of other DMARDs or use of corticosteroids once patients were included in the study, consistent with a pragmatic approach reflecting routine care.

The effectiveness results demonstrated no statistical difference between the etanercept and adalimumab groups in changes from baseline to 3 and 12 months in DAS28, CRP, patient global assessment, or tender and swollen joint count scores. The proportions of EULAR good, moderate and non-responders at 52 weeks were also assessed: these were 26.3%, 33.3% and 40.4% respectively for adalimumab versus 16.7%, 31.7% and 51.7% for etanercept (p=0.158). Adalimumab led to statistically significantly greater improvement in EQ-5D utility scores than etanercept after 52 weeks of treatment (p=0.046).

In summary, real life data from pragmatic studies and registries do not support superior efficacy of other anti-TNF agents above adalimumab. The DANBIO, DREAM and Red Sea studies, where the last was performed in a UK population eligible for biologics according to the current NICE guidance, didn't show etanercept to have superior effectiveness to adalimumab, despite being a real world mix of combination and monotherapy.

1.4 Effects of including Asian studies on MTC's effectiveness estimates

The inclusion of studies conducted in Asian populations may be a source of distortion in relative effectiveness estimates of biologics. AbbVie advocates their exclusion in the base case.

It is noted that the MTC performed by the Assessment Group did not exclude studies conducted solely in Asia – these were the cases of references number 72, 81, 89, 105 and 106 (with reference 127 used in sensitivity analysis) of the AG Report. Due to the choice of primary assessment of EULAR response by the AG and consequential limitation in number of included studies, there is a relatively high proportion of primary Asian studies for tocilizumab (TCZ) compared with the other biologics.

In AbbVie's submission, 7 studies were identified that were conducted exclusively in Asian countries (2 for adalimumab, 3 for tocilizumab, 1 for certolizumab, and 1 for golimumab). These were excluded from our base case with the rationale that the Japanese RA population differs considerably to Caucasian populations in several ways. Firstly, genetic and environmental differences have been previously noted[26]. In addition, there are significant demographic dissimilarities (including mean weight), as well as clinical differences in the treatment of RA. A key example of the latter is the dose of MTX, which was restricted to 8mg/week by Japanese treatment guidelines and the MTX package insert due to concerns of toxicity in this population[27], which could conceivably have a significant impact on clinical response in the studies.

Although studies from Asia were excluded from our base case, we performed a sensitivity analysis to assess the effect of inclusion. This had a major impact on tocilizumab monotherapy posterior mean and median ACR responses. Posterior median estimates increased by 9-10 percentage points for ACR20 and ACR50, to a level that would indicate 5 percentage point higher median response probabilities than same level ACR responses for tocilizumab in combination with MTX/DMARD. ACR estimates with Asian studies for adalimumab increased slightly, while other drugs remained unaffected. The wide variation in ACR response levels with respect to the inclusion or otherwise of Asian studies supports our assertion that studies

conducted exclusively in Asia should be omitted from the MTC when looking to make decisions in the UK.

1.5 Effects of including TACIT trial data

The inclusion of a trial – TACIT – with data marked academic in confidence compromises the transparency of the MTC, and may be a source of bias in cost-effectiveness estimates.

AbbVie is unclear regarding the rationale behind the use of the TACIT trial data[28]. This is an unpublished, open-label study comparing three anti-TNF agents as a group (in combination with DMARD) against combination cDMARDs therapy. No other similar studies appear to have been included in this analysis. As previously noted, AbbVie considers that open label studies should not be mixed with double blinded studies in the base case estimates of the MTC. It is also unclear whether the data from the cDMARDs arm has been used to inform the comparator analysis for the AG's MTC. As we do not have access to the source data, it is difficult to estimate the effect the inclusion of this trial has on the validity of the analysis. But concerns should be raised towards the inclusion of a trial that is open-label, the grouping of the anti-TNF agents together, which blurs differences in effectiveness between them, and the possible liberal use of intra-muscular steroids in the combination DMARD arm of the trial. All these could be sources of bias in the effectiveness estimates of different treatments. Long term high dose corticosteroid use is unlikely to be acceptable in clinical practice.

We would therefore request the provision of the TACIT trial data order to assess the impact of the evidence presented within the TACIT trial and its place within the evidence network as a whole. It does not seem reasonable to expect an informed critique of the methodology employed in the systematic review and subsequent meta-analysis conducted by The AG without full transparency around the included studies and the data these supply to the evidence network.

1.6 Trials not included in Systematic Review

AbbVie wishes to bring your attention to an AbbVie sponsored study which has been published as a conference poster since the original submission of evidence to NICE. The results of the study bring forward additional evidence which supports the value of adalimumab used in combination with MTX.

The CONCERTO study[29] was a 26 week, double-blind, parallel-arm trial in MTX- and biologic-naive patients with active RA of less than 1 year duration. Patients received open-label adalimumab (ADA) 40 mg every other week, and were randomised to weekly oral MTX at 2.5, 5, 10, or 20 mg.

Endpoints: The primary endpoint was the achievement of low disease activity (LDA, DAS28-CRP <3.2) at week 26. Secondary endpoints included ACR50/70/90, remission (assessed by DAS28-CRP <2.6, SDAI ≤3.3, and CDAI ≤2.8), radiographic non-progression (ΔmTSS ≤0.5), and ΔHAQ-DI ≤- 0.22 at week 26. Dose-response was assessed via Cochran-Armitage linear trend test. Non-responder imputation was used for missing data.

Results: There were 395 randomised patients, of whom 358 (83, 93, 93, and 89 from the 2.5, 5, 10, and 20 mg arms, respectively) completed 26 weeks. At week 26, there was a statistically significant trend (which can be observed in the table below) in the proportion of patients achieving LDA with increasing dose of MTX in combination with open-label ADA.

Similar results were observed for the ACR50/70/90 responses and for DAS28-CRP/SDAI/CDAI remission. Numerically increasing trends were also apparent in the proportions of patients achieving radiographic non-progression and ΔHAQDI ≤-0.22 at week 26. Clinical and radiographic outcomes generally appeared comparable between the 10 and 20 mg arms. ADA serum concentrations were similar in the 10 and 20 mg MTX arms, with slightly lower mean ADA serum concentrations in the 2.5 and 5 mg arms. Adverse events were consistent with the

known profile of ADA and were generally consistent between arms, occurring most frequently in the 20 mg arm (reproduced as Table 2 below).

Table 2: CONCERTO trial results

Parameter [a]	2.5mg MTX + OL ADA	5mg MTX + OL ADA	10mg MTX + OL ADA	20mg MTX + OL ADA	*P* value
DAS28 (CRP) <3.2, n/N (%)	42/98 (43)	44/100 (44)	56/99 (57)	59/98 (60)	0.005
DAS28 (CRP) <2.6, n/N (%)	27/98 (28)	32/100 (32)	37/99 (37)	44/98 (45)	0.008
SDAI ≤3.3, n/N (%)	11/98 (11)	22/100 (22)	28/99 (28)	29/98 (30)	0.005
CDAI ≤2.8, n/N (%)	12/98 (12)	22/100 (22)	29/99 (29)	27/98 (28)	0.02
ACR50/70/90 %	46/24/7	51/34/10	54/41/17	62/46/16	0.02/0.003/0.04
∆mTSS ≤0.5, n/N (%)	63/98 (64)	72/100 (72)	76/99 (77)	76/98 (78)	0.06
∆HAQ-DI ≤ -0.22, n/N (%)	67/98 (68)	70/100 (70)	73/99 (74)	76/98 (78)	0.12

[a] Non-responder imputation

Conclusions: MTX and biologic-naïve patients with early RA demonstrated robust and dose-dependent clinical responses with increasing doses of MTX in combination with open-label ADA through 26 weeks. Statistically significant increasing trends with increasing dose of MTX (from 2.5 mg to 20 mg) were observed for patients achieving:

- DAS28 low disease activity and remission
- SDAI low disease activity and remission
- Comprehensive disease control

Clinical relevance: Data from biologic registries suggest that approximately 1/3rd of patients take biologic drugs in monotherapy[30]. This study is particularly relevant to the RA population who are labelled 'MTX-intolerant' as it shows that robust clinical responses can be achieved even with low dose MTX.

1.7 Effectiveness measures in tocilizumab trials

AbbVie points out that the choice of response measures based on DAS28 may overestimate the effectiveness of tocilizumab, and that this will be exacerbated by choosing EULAR response over ACR response, as the weight of the acute phase response is higher in the former.

The AG report states that:

(Page 89) *"Population 2/3 […] Statistically significant findings were also identified in the ADACTA trial, whereby greater proportions of patients receiving tocilizumab monotherapy achieved ACR responses than among patients receiving adalimumab monotherapy"*

(Page 105) "*Population 2/3 […] Finally, similar reductions in swollen joint count and tender joint count were reported for patients in the tocilizumab plus placebo adalimumab and adalimumab plus placebo tocilizumab arms in the double-dummy trial ADACTA"*

We feel it is important to note that the use of any RA disease activity measure incorporating acute phase reactants (ESR or CRP) may inherently favour tocilizumab (TCZ) above biologics that do not directly inhibit the IL-6 pathway. A major physiological role of IL-6 is activation of the acute phase response including induction of CRP production by the liver[31]. There is thus potential for IL-6 inhibition to result in over-estimation of the magnitude of response to treatment, with suppression of the ESR and CRP despite persistent synovitis. This may be

particularly prominent for the DAS28-ESR and -CRP due to the high weighting of the acute phase reactant within the calculation[32], and therefore for the EULAR responses (reproduced in the Table 3 below) which are based on this outcome.

Table 3: Reproduction of EULAR responses

	DAS28 improvement		
Present DAS28	**>1.2**	**0.6-1.2**	**<0.6**
<3.2	Good response	Moderate response	No response
3.2-5.1	Moderate response	Moderate response	No response
>5.1	Moderate response	No response	No response

This issue was examined in a post-hoc analysis of the TCZ clinical trials by Smolen and colleagues[33]. Among TCZ-treated patients in DAS28-defined remission, up to 80% did not have disease remission according to the clinical disease activity index (CDAI) or simplified disease activity index (SDAI), compared with 50% of patients on Infliximab. It was concluded that

"although tocilizumab and infliximab appeared to convey numerically similar SDAI and CDAI remission rates, the differences between DAS28 remission in these two groups suggested that tocilizumab might particularly increase remission rates as assessed by the DAS28. In other words, the very low mean ESR in patients with CDAI-defined moderate disease activity "compensated" for higher values for "alternative" clinical disease activity variables (such as reflected by a higher swollen joint count)."

This was shown to be applicable to routine clinical practice in a study by Strangfeld and colleagues[34] examining the contribution of DAS28 score variables to treatment response in TCZ treated patients within the German biologics register RABBIT. In a high proportion of tocilizumab-treated patients the reduction in ESR was sufficient to reach EULAR response (nearly 35% of patients receiving TCZ compared with <5% for anti-TNF treated patients).

This issue is reflected in the Consensus Statement on IL6 inhibition developed by an international group of experts [32] which notes that;

"the effect of IL-6 inhibition on CRP levels or ESR may be profound despite lack of clinical improvement; the actual response may be obscured. Therefore measures that do not comprise an acute phase response, such as the CDAI, are preferred".

While ACR responses do include CRP or ESR in the calculation, the acute phase reactant makes the lowest contribution to the score[11]. Therefore the effect of independent acute phase reactant reduction with anti-IL6 therapy may not be as great in magnitude as it is the case with EULAR responses. No studies have examined this specific issue, to the best of AbbVie's knowledge.

This supports the assertion that if response measures which bear heavily on acute phase reactant, such as EULAR, are used, one may observe differences between biologic treatments that would not be revealed if response measures more closely linked to radiographic progression and function were chosen.

2. Issues with The AG's Account of the AbbVie Economic Model

2.1 Criticisms raised on estimation of utility equation (p. 300)

AbbVie would like to point out that due to a referencing error in the final version of the AbbVie submission the method used for HAQ to utility mapping in the base case was incorrectly identified in this version of the document. To clarify: the AbbVie base case employs the HAQ to utility mapping method described by Malottki and colleagues[35] (EQ-5D=0.804-[0.203*HAQ]-[0.045* HAQ]^2) in the base case and explores the use of the Ducournau[36] mapping method in a sensitivity analysis.

2.2 Choice of ACR50 response criteria rather than ACR20

The choice of ACR50 response criteria in the base case over the ACR20 was based on scientific considerations and taken before results using the two methods could be compared. It was also the measure used by Abbott in the previous submission of evidence for biologic treatments after cDMARDs failure.

The AG report states (p. 370):

"The minimal response required for continuation of treatment after the initial 6 month period is ACR50. The Assessment Group note that the comparative results for AbbVie's intervention (adalimumab) appears to perform relatively better using ACR50 than by using ACR20"

Choice of ACR50 as the goal of biologic therapy for continuation of therapy was chosen based on recent debates [37] on whether patients achieving only 20% response should be switched to another biologic agent to attempt to achieve this greater improvement.

The AG suggests that results for adalimumab were better using ACR50 rather than ACR20 as the cut-off for continuation of biologics. Use of the ACR50 instead of ACR20 was however unfavourable to adalimumab monotherapy in all populations evaluated and unfavourable to adalimumab combination therapy amongst severe methotrexate naive patients. The results in our submission support AbbVie's view. The ICER versus cDMARDs was slightly more favourable (5% reduction) for adalimumab combination therapy when comparing amongst moderate and severe MTX experienced patients; however it should be noted that in these two populations the use of ACR50 resulted in combination therapy with adalimumab being extendedly dominated when assessing incremental results. Use of ACR20 as a cut-off however resulted in combination therapy with adalimumab being cost-effective versus DMARDs but also being on the frontier, thus AbbVie disagrees that use of the ACR50 provided better results for adalimumab.

2.3 Correlation parameters in AbbVie's estimates of treatment discontinuation

The AG correctly states that (p.280 of the report)

"The fitted curves to the data are shown in Table 134. The correlation between the parameters was not provided in the report."

A table with the correlation parameters column added is reproduced below as Table 4.

Table 4: Reproduction of Table 134 from p.280 of AG report

Treatment	Lambda		Gamma		Rho
	Mean	SE	Mean	SE	
MTX	2.1163	0.0531	2.8986	0.0472	0.001299
MTX+HCQ[a]	2.1163	0.0531	2.8986	0.0472	0.001299
SSZ+HCQ[a]	2.1163	0.0531	2.8986	0.0472	0.001299
LEF[a]	2.1163	0.0531	2.8986	0.0472	0.001299
HCQ	0.4165	0.0802	2.1706	0.0674	0.001213
SSZ	0.6336	0.0303	2.4548	0.0259	0.000204
CYC[b]	0.6336	0.0303	2.4548	0.0259	0.000204

CYC=ciclosporin; HCQ=hydroxychloroquine; LEF=leflunomide; MTX=methotrexate; SE=standard error; SSZ = sulfasalazine; a=Assume similar time on treatment as methotrexate; b=Assume similar time on treatment as sulfasalazine

2.4 Modelling of survival using Gompertz curves

The AG correctly states (p. 352 of the report):

"The Assessment Group comment that no goodness of fit values for the Gompertz model compared with the life table data were presented."

AbbVie comments that the Gompertz distribution was used as there is evidence to suggest it is a good descriptor for all causes of death, while for example the Weibull distribution has shown to be a better descriptor of purer, single causes of death[38].

3. Issues with the AG's Model

3.1 Modelling of HAQ progression

3.1.1 Data sources used to model HAQ progression

AbbVie considers that the populations whose data were used to model the progression of HAQ scores while on cDMARDs do not reflect the populations to which this assessment should apply. The anticipated lower HAQ progression in the populations adopted is likely to have resulted in an overestimation of ICER values for biologics with respect to cDMARDs.

We question the use of the ERAS cohort [39] to calculate the trajectory of RA patients treated with cDMARDs. The ERAS cohort included patients with relatively early disease and before biologics therapy was commenced, and patients were recruited between 1986 and 1998. The ERAS cohort included all patients with suspected early RA, which would include patients with mild and even self-limiting disease. Patients with mild RA are not eligible for treatment with biologic agents in the UK and therefore this cohort could conceivably underestimate HAQ progression for the select group of patients who should be eligible for biologic therapy (DAS >3.2 who have failed combination DMARD therapy).

A more realistic cohort to determine trajectory of HAQ progression on cDMARD therapy, at least for populations 2 and 3, would be a more recently recruited patient group with active RA despite DMARD therapy, such as the BSRBR comparator (cDMARDs) cohort. These patients have chronic severe RA, which is more reflective of the spectrum of patients who would be treated with biologics in the UK.

AbbVie notes that, even if The AG validated the HAQ progression estimation based on ERAS data with different datasets (based on the Norfolk Arthritis Register patients and the Early Rheumatoid Arthritis Network), the ERAN cohort[40], which recruited patients from 2002, even if it would be more reflective of modern treatment practices, still included patients with mild RA, which would be ineligible for biologics therapy. Neither of these populations would be suitable for patients presenting with new DAS28 above 3.2, as early treatment of these patients to a pre-defined target is key, as reflected by Best Practice Tariffs issued by the Department of Health and NICE's Quality Standards for Rheumatoid Arthritis.

This makes it likely that average HAQ progression in the ERAS cohort will have been slower in ERAS patients than it would in a cohort of patients eligible for biologics. This may be the source of underestimation of HAQ progression on cDMARDS, which can result in ICERs estimates of biologics with respect to cDMARDs being higher.

3.1.2 Lack of transparency in AG's modelling of HAQ progression

AbbVie notes that the modelling on HAQ progression in the AG's model is not transparent, and that the modelling choices have large effects on cost-effectiveness estimates whose mechanism is hard to discern.

The choice of patient population from which HAQ progression on cDMARDs was estimated has been criticized in Section 3.1.1. AbbVie further points out that The AG's model is subject to a large degree of parameter uncertainty, as the cost-effectiveness results are also highly sensitive to the methods used to model HAQ progression on cDMARDs and biologics. Large drops in ICERs of biologics with respect to cDMARDs are observed if the linear HAQ progression assumptions used in previous submissions (detailed in the second paragraph in p. 365 of the AG report), which were used in scenario analysis by The AG, had been adopted.

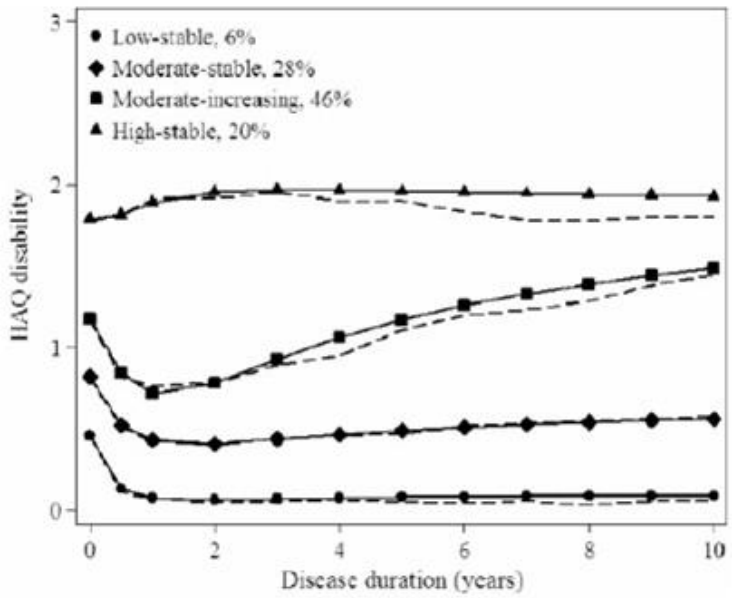
Not having had access to The AG's model, AbbVie is unable to comment on the implementation of the HAQ trajectories on different classes of patients and how they are defined. In its modelling of the progression of HAQ following initial response while on cDMARDs, The AG applied a Latent Class Growth Model to data from the Early Rheumatoid

Arthritis Study (ERAS), presenting (after request from AbbVie) a graphical representation of HAQ progression over time for the four different classes of patients identified by the model. While this provides additional insight into the modelling of HAQ progression, it should be noted that figure 3 in reference 275 does not depict the same behaviour of patient classes as the second panel in the figure in reference 270 (both are reproduced below in this order), with the first figure showing a remarkably less pronounced HAQ progression across various classes. It is not clear what justifies the differences, as both studies were conducted on the ERAS database. This also puts into question the validity of the assertion by the AG that (p.371 of the report):

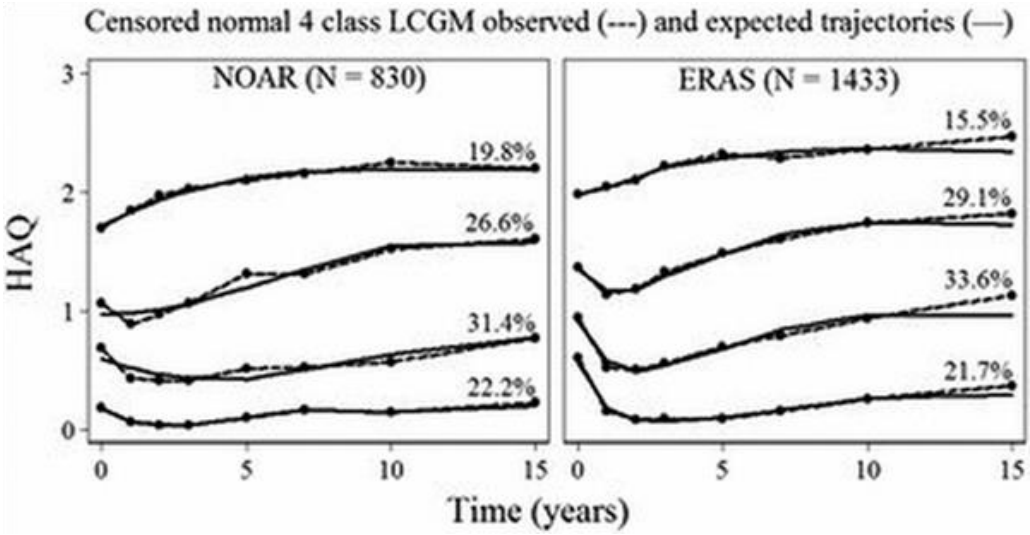
"In all cases, the rate of worsening over time decreases. This is directly contrary to the typical assumptions of DMARD worsening incorporated into cost effectiveness models."

as this does not look the case for the Moderate-Increasing group in figure 3 in reference 275, which represents 46% of the patient population.

FIG. 3 HAQ disability trajectories for the four-class GMM.



Solid lines are estimated trajectories and dashed lines are observed trajectories for each class.



The model does not allow consultees to discern how HAQ scores would progress in the average cohort in the model. This precludes comparisons with estimates from literature, obscures the effect that HAQ progression modelling has on cost-effectiveness results, and does not allow AbbVie to incorporate what would be the AG's estimates of HAQ progression for the average cohort in its model. It is therefore not feasible to check if shifts in ICERs similar to the ones observed in The AG's model would arise from changing the HAQ progression estimates adopted by AbbVie in its base case to whole population average estimates of HAQ progression as estimated by The AG.

3.2 HAQ progression on cDMARDs vs. biologics

AbbVie considers that the assumption that HAQ change, conditional on EULAR response, is the same for cDMARDs and biologics is not supported by clinical evidence, and is a source of bias which ultimately will worsen the cost-effectiveness of estimates of biologics with respect to cDMARDs. This is discussed on the AG report (p. 364):

"In the base case the Assessment Group assume that the HAQ change, conditional on EULAR response, was the same for cDMARDs as for bDMARDs."

This assumption is not grounded on evidence and is a source of bias in effectiveness estimates, which disfavours biologic treatments.

Firstly, in RA patients with poor clinical response to therapy, although structural damage in RA has been correlated with volume of synovitis in patients on cDMARDs, it has been demonstrated that many of the currently licensed biologics (including adalimumab, etanercept, infliximab, rituximab, and tocilizumab) can inhibit radiographic progression even in patients who fail to respond clinically, while this is clearly not the case for MTX monotherapy[41,42,43,44,45]. To our knowledge, whether the same is true for combination DMARD therapy versus MTX monotherapy has not been examined. However this is unlikely, as the mechanism for the disconnect between radiographic progression and inflammation with biologics is thought to be at least partly attributable to a direct osteoclastic inhibition, in addition to an indirect reduction in synovitis. In contrast, the mechanism of radiographic inhibition with DMARDS including MTX is thought to predominantly relate to a reduction in the burden of inflammation.

Therefore, in the presence of on-going inflammatory activity including subclinical synovitis[46] (e.g. any response other than a good EULAR response resulting in remission at the subclinical level), treatment with biologics would likely result in lower radiographic progression compared with cDMARDs. In chronic RA, HAQ is longitudinally correlated with radiographic progression [47]

and therefore AbbVie does not believe that the HAQ change for a given EULAR response would be the same for cDMARDs as for bDMARDs, even if a patient fails to respond to treatment.

Secondly, in patients with positive clinical response to therapy, it has also been shown that structural damage can accrue despite a good clinical response to treatment. A post hoc analysis of the PREMIER study demonstrated a strong proportional relationship between clinical and radiographic efficacy for MTX monotherapy, but not for adalimumab in combination with MTX or adalimumab monotherapy across the ACR 20, 50 and 70 responses over 104 weeks [48]. This supports the notion that radiographic inhibition with MTX is dependent on reduction of synovitis, while this is not the case for biologics (particularly when used in combination with MTX). Although we are unaware of any studies assessing this relationship according to the EULAR response criteria, DAS28-defined remission was examined in this study with similar results. Indeed inhibition of radiographic progression with MTX monotherapy approximated that for adalimumab plus MTX only for patients who achieved very strict definitions of clinical response (0 swollen joints or an ACR100 response)[41,48].

Finally, it has also been shown that structural damage can continue despite ACR-defined and DAS-defined remission over 2 years for patients taking a variety of cDMARDs including sulphasalazine, gold, and MTX[49].

The importance of radiographic progression despite clinical response with cDMARDs is underpinned by the knowledge that clinical examination under-estimates the degree of synovitis as determined by US or MRI imaging, even in patients considered to be in clinical remission[46].

Again, we are not aware of data demonstrating the relative HAQ change dependent on DAS28 remission or good EULAR response in patients treated with biologics relative to those on cDMARD therapy. However, the relationship between structural damage and HAQ is assumed to apply in this situation.

It has been demonstrated that early joint damage in patients with recent onset RA is related to subsequent HAQ-disability, and this is largely due to changes in joint space narrowing rather than in erosion score[44, 50]. Adalimumab plus MTX has been shown to significantly inhibit progression of both erosions and joint space narrowing independently of disease activity, although the effect on the latter was the most striking. This supports the effectiveness of adalimumab in combination with MTX in preventing radiographic progression and improving functional outcomes.

3.3 Modelling based on EULAR response rather than ACR response

AbbVie considers that an MTC based on ACR response would be more transparent than the one based on EULAR response which was adopted in the AG's model.

It is very unclear from the Draft Appraisal Report precisely how ACR data was mapped on to EULAR response. In particular, results from the VA analysis are redacted as academic-in-confidence (AG report, Table 187, p.380) and so the face validity of the mapping cannot be established. Furthermore, it is somewhat concerning that the approach does not appear to control for factors such as age, gender, disease history and disease activity which may affect the mapping between ACR and EULAR response at the individual level (and therefore, could induce an inverse ecological bias when scaling up to trial-level response rates). Finally, even if the estimated mapping approach is valid, we have been given no indication as to its statistical precision and how/whether the uncertainty induced by this mapping has been appropriately propagated through the model. The lack of transparency is particularly worrying as there are sizeable differences in cost-effectiveness results obtained using EULAR responses and ACR responses mapped to EULAR responses which cannot be further investigated. AbbVie considers that the lack of transparency regarding this mapping, limits the ability of stakeholders to comment on the estimates of cost effectiveness obtained using these data.

3.4 Costs to society beyond direct medical costs not considered in the AG's model

AbbVie considers that the wider costs to society of RA and the benefits derived from its appropriate management should be presented, as a substantial decrease in wider societal costs can cause a sizeable drop in ICERs of biologics with respect to conventional DMARDs that should be investigated.

AbbVie included in its submission a sensitivity analysis which looked at the effect of including costs other than just direct medical costs in the model. The wider costs to society of RA are substantial. A recent report by The University of Manchester[51] reviews the evidence and quotes a report from the National Audit Office which estimated that RA healthcare costs in England amounted to £560 million per year[52], with estimates rising to £1.8 billion a year if the cost of sick leave and lost employment was included. Also quoted is a study by Young and colleagues which concluded that, in a five year follow-up of 732 early RA patients, by the end of the follow-up period approximately a fifth had retired due to their RA[53]. In addition to the impact of RA treatments on productivity and absenteeism, factors such as impact on the costs associated with caring for a family member or friend with RA and costs incurred by patients purchasing aids and devices to assist with reduced mobility caused by RA should also be considered. The relative importance of these costs with respect to direct medical costs cannot be understated, and a report where these are not mentioned gives only a partial view of the true cost-effectiveness estimates of treatments.

3.5 Costs of administration of drugs

AbbVie considers that the cost of infusion of intravenous (IV) drugs and the costs of administration of subcutaneous (SC) drugs are not correctly specified in the AG's model.

The AG report states (p. 355) that

"The administration costs of infusions were taken from TA247 in which the final appraisal determination (FAD) stated that 'the manufacturer's revised estimate of £154 was acceptable'. This estimate (of 60 minutes infusion time was also applied to abatacept and infliximab) in the absence of a robust relationship between costs and infusion times. This assumption may be favourable to infliximab and unfavourable to abatacept as the recommended infusion times are at least 2 hours, and 30 minutes respectively. The FAD for TA247 did not comment on the assumption that 10% of subcutaneous injections would be performed by district nurses and the Assessment Group has assumed that these were also thought acceptable. This resulted in an average administration cost per subcutaneous injection of £2.61."

By assuming a cost of infusion common to all IV drugs the AG has not addressed the issue of differences in administration of costs between IV drugs. Within the framework of THE AG's model, two different unit costs should be applied, corresponding to DoH tariffs which would be either Day Case or Outpatient Tariffs, depending on the duration of administration of drugs. An Outpatient tariff could be applied for shorter administration time drugs (possibly around 3 hours or less); whereas a higher Day Case rate could be applied to drugs whose administration time exceeded around 3 hours. Moving from Day Case to Outpatient Tariffs over time may be envisaged if infusion time is reduced (e.g. to account for less necessity for close monitoring of adverse events).

With respect to subcutaneous injections, AbbVie has the view that administration cost should be set to zero in the base case, as the vast majority of patients requiring training and help with SC administration of adalimumab receive this service via Homecare providers, which are fully funded by AbbVie.

3.6 Impact of HAQ scores on mortality comes from HAQ values at baseline only

AbbVie points out that this assumption may be a source of bias in the results, as it ignores the effect that a more favourable HAQ progression on patients treated with biologics with respect to patients treated with cDMARDs and may result in underestimation of additional survival provided by biologics over cDMARDs.

The AG assumed that only the baseline HAQ score was important for predicting mortality. Whilst the AG is right to point out that the study by Michaud and colleagues [54] found that;

""changes in the PCS [SF36 physical component summary score] and HAQ did not contribute substantially to predictive value over and above the baseline values of these variables""

there remain associations between changes in HAQ over time and mortality risk. For example the same study reports that an improvement in HAQ of >=0.5 units was associated with a significant reduction in mortality risk over all time-periods from 6 months through to three years. Similarly HAQ worsening was always associated with increased mortality risk. Thus whilst we agree that starting level of the PCS and HAQ are more important than the magnitude of change in HAQ, there is evidence to suggest that changes in HAQ affect mortality risk. It should be noted that the study was conducted on a much less severe cohort of patients than the BSRBR comparator cohort, so more evidence would be needed to ascertain the generalizability of these results to a more severe cohort of patients.

Based on sensitivity analysis conducted with the AbbVie model, the impact of this assumption is not large when considering the incremental cost per QALY gained; however it is important to note that since changes in HAQ do influence mortality, one would expect differences in estimates of survival in patients treated with bDMARDs versus cDMARDs. Thus this assumption is conservative towards bDMARDs and favours cDMARDs.

3.7 Incidence of adverse events is different for different biologic treatments

AbbVie considers that the assumption of equal rate of serious infection across all biological is not the best taking into account existing evidence.

The AG states (p. 381 of the report):

"It was assumed that the rate of serious infection was independent of the bDMARDS used."

AbbVie wishes to highlight that this assumption is contrary to the findings of a recent network meta-analysis included in a Cochrane review which examined the adverse effects of biologics across any indication except HIV/AIDS, including RA. Certolizumab pegol was associated with significantly higher risk of serious infections compared to control treatment in direct comparisons (Odds Ratio (OR) 4.75, 95% Confidence Interval 1.52 to 18.45) with a Number Needed to Harm of 12 (4 to 79). Indirect comparisons against other biologics demonstrated that certolizumab was associated with higher odds of serious infections compared with abatacept, adalimumab, etanercept, golimumab and rituximab; the OR were about 0.25-times or lower for each of the five biologics compared with certolizumab[55].

3.8 Cost of treating serious infections

The cost of treating serious infections may be underestimated.

The AG report states (p.382) that:

"The costs (£1479 per episode) and undiscounted QALY loss associated with serious infections (a loss in utility of 0.156 for 28 days) were both taken from the Pfizer submission"

These are specified in Table 154 (p.311 of the report), which states that

"Within the adverse events scenario analysis, the cost of serious infection was assumed to be £1,497 based on relevant NHS costs, weighted by inpatient activity. Relevant HRG codes were identified based on Lekander et al, 2010 conservatively the without complications and contraindications HRG costs were used."

The base for identifying relevant DRG codes was a paper in a Swedish setting[56]. It should be made clear if the DRG codes would be relevant in the UK setting and if the difference between the price of treating serious infections in the article (€ 3,321) and the cost of treating them in the UK would be expected to take into account different management strategies and general cost levels in both healthcare systems. Also, the impact of considering HRGs without complications and contraindications on the cost of treating adverse events should be discussed and values for less conservative assumptions presented.

References

1 Fransen J et al . Disease Activity Scores using C-reactive protein: CRP may replace ESR in the assessment of RA disease activity. Ann Rheum Dis 2004;62:151.

2 Soubrier M, Zerkak D, Gossec L, Ayral X, Roux C, Dougados M. Which variables best predict change in rheumatoid arthritis therapy in daily clinical practice? J Rheumatol. 2006 Jul;33(7):1243-6. Epub 2006 Apr 15.

3 Crowson CS, Rahman MU, Matteson EL.Which measure of inflammation to use? A comparison of erythrocyte sedimentation rate and C-reactive protein measurements from randomized clinical trials of golimumab in rheumatoid arthritis. J Rheumatol. 2009 Aug;36(8):1606-10. doi: 10.3899/jrheum.081188.

4 Matsui T et al. Disease Activity Score 28 (DAS28) using CRP underestimates the disease activity and overestimates the EULAR response criteria compared with DAS28 using ESR in a large observational cohort of rheumatoid arthritis patients in Japan. Ann Rheum Dis 2007;66:1221–6.

5 Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of Disease Activity Score (DAS)28-erythrocyte sedimentation rate and DAS28-C-reactive protein threshold values. Ann Rheum Dis 2007;66:407

6 Hensor EM, Emery P, Bingham SJ, Conaghan PG; YEAR Consortium. Rheumatology (Oxford).Discrepancies in categorizing rheumatoid arthritis patients by DAS-28(ESR) and DAS-28(CRP): can they be reduced? 2010 Aug;49(8):1521-9. doi: 10.1093/rheumatology/keq117. Epub 2010 Apr 30.

7 Castrejón I, Ortiz AM, Toledano E, Castañeda S, García-Vadillo A, Patiño E, González-Alvaro I.Estimated cutoff points for the 28-joint disease activity score based on C-reactive protein in a longitudinal register of early arthritis. J Rheumatol. 2010 Jul;37(7):1439-43. doi: 10.3899/jrheum.091333. Epub 2010 May 15.

8 Tamhane A et al. Comparison of the Disease Activity Score Using Erythrocyte Sedimentation Rate and C-reactive Protein in African Americans with Rheumatoid Arthritis. J Rheumatol. 2013 Aug 15.

9 Wells G et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009;68:954-960 doi:10.1136/ard.2007.084459.

10 Tamhane A et al. Comparison of the Disease Activity Score Using Erythrocyte Sedimentation Rate and C-reactive Protein in African Americans with Rheumatoid Arthritis. J Rheumatol. 2013 Aug 15.

11 Paulus H. Relative contributions of the components of the American College of Rheumatology 20% criteria for improvement to responder status in patients with early seropositive rheumatoid arthritis. Arthritis Rheum. 2000 Dec;43(12):2743-50.

12 Mills EJ, Ghement I, O'Regan C, Thorlund K. Estimating the Power of Indirect Comparisons: A Simulation Study. Plos One. 2011 Jan 21;6(1):e16237.

13 Soares MO, Dumville JC, Ades AE and Welton NJ. Treatment comparisons for decision making: facing the problems of sparse and few data. J. R. Statist. Soc. A (In Press. 2014)

14 Keystone, E.C., Wang, M.M., Layton, M., Hollis, S., McInnes, I.B., Study Team. Clinical evaluation of the efficacy of the P2X7 purinergic receptor antagonist AZD9056 on the signs and symptoms of rheumatoid arthritis in patients with active disease despite treatment with methotrexate or sulphasalazine.[Erratum appears in Ann Rheum Dis. 2012 Dec;71(12):2064]. *Annals of the Rheumatic Diseases* 2012; 71(10):1630-1635.

15 Combination Etanercept and Methotrexate Therapy Provides Better Outcomes Than Standard DMARD and Methotrexate Therapy in Rheumatoid Arthritis: Results From a Study in the Latin America Region. 2012.

16 van Vollenhoven, R.F., Kinnman, N., Vincent, E., Wax, S., Bathon, J. Atacicept in patients with rheumatoid arthritis and an inadequate response to methotrexate: Results of a phase II, randomized, placebo-controlled trial. *Arthritis and Rheumatism* 2011; 63(7):1782-1792.

17 Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. Stat Med 2007;26(6):1237-54.

18 O'Dell JR et al. Therapies for active rheumatoid arthritis after methotrexate failure. N Engl J Med. 2013 Jul 25;369(4):307-18. doi: 10.1056/NEJMoa1303006. Epub 2013 Jun 11.

19 NCCfC. National Clinical Guideline for Rheumatoid Arthritis. Royal College of Physicians; 2008.

20 Karlsson et al. Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in early rheumatoid arthritis: 2-year quality-of-life results of the randomised, controlled, SWEFOT trial. Ann Rheum Dis. 2012 Nov 29. [Epub ahead of print]
21 Moreland LW et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. Arthritis Rheum. 2012 Sep;64(9):2824-35. doi: 10.1002/art.34498
22 van Vollenhoven RF, Cifaldi MA, Ray S, Chen N, Weisman MH. Improvement in work place and household productivity for patients with early rheumatoid arthritis treated with adalimumab plus methotrexate: work outcomes and their correlations with clinical and radiographic measures from a randomized controlled trial companion study. Arthritis Care Res (Hoboken). 2010 Feb;62(2):226-34. doi: 10.1002/acr.20072.
23 Hetland, M.L., Christensen, I.J., Tarp, U., Dreyer, L., Hansen, A., Hansen, I.T. et al. Direct comparison of treatment responses, remission rates, and drug adherence in patients with rheumatoid arthritis treated with adalimumab, etanercept, or infliximab: results from eight years of surveillance of clinical practice in the nationwide Danish DANBIO registry. Arthritis & Rheumatism 2010; 62(1):22-32.
24 Kievit W, Adang EM, Fransen J, Kuper HH, van de Laar MA, Jansen TL, et al. The effectiveness and medication costs of three anti-tumour necrosis factor α agents in the treatment of rheumatoid arthritis from prospective clinical practice data. Ann Rheum Dis 2008; 67: 1229–34
25 Jobanputra, P., Maggs, F., Deeming, A., Carruthers, D., Rankin, E., Jordan, A.C. et al. A randomised efficacy and discontinuation study of etanercept versus adalimumab (RED SEA) for rheumatoid arthritis: a pragmatic, unblinded, non-inferiority study of first TNF inhibitor use: outcomes over 2 years. *Bmj Open* 2012; 2(6).
26 Turkstra E, Ng S-K, Scuffham PA. A mixed treatment comparison of the short-term efficacy of biologic disease modifying anti-rheumatic drugs in established rheumatoid arthritis. Current Medical Research and Opinion. 2011 Oct;27(10):1885–97.
27 Miyasaka N. Treatment trends of rheumatoid arthritis in Japan: changes toward globalization and its unique innovation. Inflamm Regen 2011;31:25–32
28 Randomised controlled trial of tumour-necrosis-factor inhibitors against combination intensive therapy with conventional disease modifying anti rheumatic drugs in established rheumatoid arthritis: the TACIT trial http://www.controlled-trials.com/ISRCTN37438295
29 Burmester et al. Efficacy, pharmacokinetics, and safety of different doses of methotrexate in combination with adalimumab: results from the CONCERTO trial. Presented at EULAR 2013.
30 Gabay C et al. Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. Lancet. 2013 May 4;381(9877):1541-50. doi: 10.1016/S0140-6736(13)60250-0. Epub 2013 Mar 18.
31 Nishimoto et al. Interleukin 6: from bench to bedside. Nat Clin Pract Rheumatol 2006;2:619-26.
32 Smolen JS et al. Consensus statement on blocking the effects of interleukin-6 and in particular by interleukin-6 receptor inhibition in rheumatoid arthritis and other inflammatory conditions. Ann Rheum Dis. 2013 Apr;72(4):482-92. doi: 10.1136/annrheumdis-2012-202469. Epub 2012 Nov 21
33 Smolen JS, Aletaha D. Interleukin-6 receptor inhibition with tocilizumab and attainment of disease remission in rheumatoid arthritis: the role of acute-phase reactants. Arthritis Rheum. 2011 Jan;63(1):43-52. doi: 10.1002/art.27740.
34 Strangfeld, et.al. Am College of Rheumatology 2010, Abstr #1112 http://www.biologikaregister.de/pnadmin/contents/ergebnisse_pdf/RABBIT_Tocilizumab.pdf
35 Malottki K et al. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation. Health Technol Assess. 2011 Mar;15(14):1-278. doi: 10.3310/hta15140.
36 Ducornau P, Kielhorn A, Wintfeld N. Comparison of linear and non-linear utility mapping between HAQ and EQ-5D using pooled data from the tolicizumab trials OPTION and LITHE. Annual Meeting of the British Society of Rheumatology (BSR), Glasgow 2009;(Poster 258).
37 Chung CP, Thompson JL, Koch GG, Amara I, Strand V, Pincus T. Are American College of Rheumatology 50% response criteria superior to 20% criteria in distinguishing active aggressive treatment in rheumatoid arthritis clinical trials reported since 1997? A meta-analysis of discriminant capacities. Ann Rheum Dis. Dec 2006;65(12):1602-1607.

38 Juckett DA, Rosenberg B. Comparison of the Gompertz and Weibull functions as descriptors for human mortality distributions and their intersections. Mech Ageing Dev. 1993 Jun;69(1-2):1-31.
39 Young A, Dixey J, Cox N et al. How does functional disability in early rheumatoid arthritis (RA) affect patients and their lives? Results of 5yr follow up in 781 patients from the Early RA Study (ERAS). Rheumatology 2000;39:603–11
40 Kiely P, Williams R, Walsh D, Young A: Contemporary patterns of care and disease activity outcome in early rheumatoid arthritis: the ERAN cohort. Rheumatology (Oxford) 2009, 48(1):57-60.
41 Aletaha D, Smolen JS. Joint damage in rheumatoid arthritis progresses in remission according to the Disease Activity Score in 28 joints and is driven by residual swollen joints. Arthritis Rheum. 2011 Dec;63(12):3702-11. doi: 10.1002/art.30634.
42 Landewe R, van der Heijde D, Klareskog L, et al. Disconnect between inflammation and joint destruction after treatment with etanercept plus methotrexate. Arthritis Rheum 2006;54:3119–25.
43 Smolen JS et al. Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study. Arthritis Rheum. 2005 Apr;52(4):1020-30.
44 Smolen JS, van der Heijde DM, Keystone EC, et al. Association of joint space narrowing with impairment of physical function and work ability in patients with early rheumatoid arthritis: protection beyond disease control by adalimumab plus methotrexate. Ann Rheum Dis 2012. Published Online First 22 Aug2012.
45 Smolen JS, Martinez Avila JC, Aletaha D. Tocilizumab inhibits progression of joint damage in rheumatoid arthritis irrespective of its anti-inflammatory effects: disassociation of the link between inflammation and destruction. Ann Rheum Dis 2012;71:687–93.
46 Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, Wakefield RJ, O'Connor PJ, Emery P. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum. 2006 Dec;54(12):3761-73.
47 Ødegård S, Landewé R, van der Heijde D, Kvien TK, Mowinckel P, Uhlig T. Association of early radiographic damage with impaired physical function in rheumatoid arthritis: a ten-year, longitudinal observational study in 238 patients. Arthritis Rheum. 2006 Jan;54(1):68-75.
48 Emery P, Genovese MC, van Vollenhoven R, Sharp JT, Patra K, Sasso EH. Less radiographic progression with adalimumab plus methotrexate versus methotrexate monotherapy across the spectrum of clinical response in early rheumatoid arthritis. J Rheumatol. 2009 Jul;36(7):1429-41. doi: 10.3899/jrheum.081018. Epub 2009 Apr 15.
49 Molenaar ET, Voskuyl AE, Dinant HJ, Bezemer PD, Boers M, Dijkmans BA.Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. Arthritis Rheum. 2004 Jan;50(1):36-42.
50 Maillefert JF, Combe B, Goupille P, Cantagrel A, Dougados M. The 5-yr HAQ-disability is related to the first year's changes in the narrowing, rather than erosion score in patients with recent-onset rheumatoid arthritis. Rheumatology (Oxford). 2004 Jan;43(1):79-84. Epub 2003 Nov 17.
51 Sarah Parsons S, Ingram M, Clarke-Cornwell AM and Symmons DMP. A Heavy Burden: The Occurrence and Impact of Musculoskeletal Conditions in the United Kingdom Today. Available at http://www.inflammation-repair.manchester.ac.uk/musculoskeletal/aboutus/publications/heavyburden.pdf (accessed 17 September 2013).
52 National Audit Office, Morse A. Services for people with rheumatoid arthritis. London: Stationery Office, 2009.
53 Young A, Dixey J, Cox N, Davies P, Devlin J, Emery P et al. How does functional disability in early rheumatoid arthritis (RA) affect patients and their lives? Results of 5 years of follow-up in 732 patients from the Early RA Study (ERAS). Rheumatology.(Oxford) 2000;39:603–11.
54 Michaud, K., Vera-Llonch, M., Oster, G. Mortality Risk by Functional Status and Health-related Quality of Life in Patients with Rheumatoid Arthritis. Journal of Rheumatology 2012; 39(1):54-59.
55 Singh JA, Wells GA, Christensen R, Tanjong Ghogomu E, Maxwell L, et al. Adverse effects of biologics: a network meta-analysis and Cochrane overview. Cochrane Database Systematic Review 2011 Feb 16;(2):CD008794.

56 Lekander I, Borgström F, Lysholm J, van Vollenhoven RF, Lindblad S, Geborek P, Kobelt G. The cost-effectiveness of TNF-inhibitors for the treatment of rheumatoid arthritis in Swedish clinical practice. Eur J Health Econ. 2012 Sep 19. [Epub ahead of print]