23rd September 2013

Bristol-Myers Squibb
[redacted]
[redacted]

National Institute for Health and Care Excellence (NICE)
Lori Farrar
Technology Appraisal Project Manager – Committee C
Centre for Health Technology Evaluation
Level 1a, City Tower
Piccadilly Plaza
Manchester
M1 4BT

Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, abatacept and tocilizumab for the treatment of rheumatoid arthritis (review of TA guidance 130, 186, 224, 234 and part review of TA guidance 225 and 247) [ID537] – Comments on Assessment Group Report.

Dear Lori,

Thank you for your invitation to comment on the Assessment Group report for the ongoing NICE MTA for Rheumatoid Arthritis.

Firstly, we would like to highlight that previous NICE assessments of the aforementioned treatment options for Rheumatoid Arthritis (RA) have concluded that these treatment options are a cost-effective use of NHS resources and were subsequently given positive HTA recommendations. Very little (if anything at all) has changed since the publication of these HTAs, in terms of the clinical evidence base, clinical practice/management of RA in the NHS, and the cost of these treatment options.

However, the Assessment Group report for the ongoing NICE RA appraisal (ID537) now indicates that no biologic therapies for the treatment of RA can be considered cost-effective as 'The Assessment Group believes the ICERs for bDMARDs....is credibly greater than £50,000 per QALY when compared to a cDMARD alone strategy'. Whilst we appreciate that the results of the previous economic evaluations (TA130, 186, 224, 234, 225 and 247) varied, the majority of accepted ICERs were less than £30,000 per QALY, and cost-effective conclusions were reached.

AG model structure and assumptions

We are concerned that the Assessment Group model structure and assumptions do not replicate NICE guidance on biologic drugs for the treatment of rheumatoid arthritis. Our main concerns with the model structure are summarised below:

- The AG model assumes that a good or moderate EULAR response is the same as a DAS28 improvement of ≥1.2 – this is not accurate because a patient can be classified as a moderate responder whilst achieving a DAS28 improvement between 0.6 and 1.2 (if their DAS28 at endpoint is less than or equal to 5.1)
- The AG model assumes that if a biologic drug is withdrawn because of an AE within the first 6 months, then the patient receives rituximab – whereas NICE guidance states an alternative biologic should be considered
- The AG model assumes that after their first biologic, all patients receive rituximab followed by tocilizumab – and NICE guidance does not specifically state this

Additionally, BMS believe there are a number of other assumptions within the AG model which may not be appropriate:

- Assuming that HAQ improvement conditional on EULAR response is the same for all treatments could underestimate the effectiveness and hence cost-effectiveness of bDMARDs
- The HAQ trajectory analysis changes the AG model results by nearly 100%, but only changes the BMS model results by 30-40%
- Assuming treatment effect on pain is directly proportional to effect on HAQ may not be appropriate and may double-count the HAQ benefit

Patient population

We do not believe that our interpretation of the populations in the scope were incorrect as stated in the AG report. The populations were not specifically defined in the scope and were left to the interpretation of the company. We do not believe that the approach taken by the AG, to apply efficacy data from a mixed population of moderate and severe patients equally to moderate only and severe only patients, is appropriate.

Interpretation and implication of results

The Assessment Group report states that the ICERs 'are marginally higher for moderate RA patients'. We understand how this happens, however, in clinical practice this is illogical. It is widely recognised that early therapeutic intervention, compared to late intervention, is likely to improve long term outcomes, prevent joint damage, and limit functional decline, helping people maintain their mobility and independence and reduce work related disability. We would therefore expect the ICERs for moderate RA patients to be lower than those for severe disease and this leads us to question why the Assessment Group has employed a methodology that implies otherwise.

We accept that the ICERs for bDMARDs v cDMARDs are useful for comparing the cost-effectiveness between the bDMARDS, however, we believe they cannot be used to evaluate the cost-effectiveness of first-line bDMARDs compared to cDMARDs alone.

We encourage NICE to consider the implication of the AG model assumptions within treatment sequences beyond first-line biologics. BMS believe that if the cost-effectiveness of bDMARDs preceding an existing sequence of biologic treatments is to be compared to cDMARDs alone, the cost-effectiveness of the treatment sequence following first bDMARD should be considered. The treatment sequence following first bDMARD has not been demonstrated to be cost-effective using the AG model.

BMS strongly request that NICE consider the implication of either not recommending any of the bDMARDs. If first-line bDMARDs are not recommended as they are not deemed cost-effective, there will be implications for subsequent lines of treatment which have only been recommended (or appraised) with bDMARDs as first-line.

BMS are concerned that this Assessment Group report indicates that no biologic therapies for the treatment of RA can be considered cost-effective, when previous NICE assessments, in the not too distant past, of these treatment options concluded that they are a cost-effective use of NHS resources; especially given the lack of changes in the clinical evidence base, and the cost of these treatment options.

All of the above key points are discussed in further detail in our technical review below.

We would greatly appreciate it if you would consider the points that we make in this response prior to the meeting of the Appraisal Committee on Tuesday 15th October.

If you have any questions please do not hesitate to contact me.

Yours sincerely,

[redacted]

[redacted], [redacted]
Bristol-Myers Squibb Pharmaceuticals Limited

Bristol-Myers Squibb

Technical Review

Assessment Group model structure

The AG model does not replicate NICE guidance on biologic drugs for the treatment of rheumatoid arthritis and so the results it generates are not necessarily applicable to current clinical practice.

The Assessment Report states that:

> "None of the models submitted by the manufacturers replicated the clinical reality within England and Wales to the satisfaction of the Assessment Group." – Section 6.3, page 346

However, BMS do not believe that the AG model reflects clinical reality within England and Wales, for the following reasons:

1. The AG model assumes that a good or moderate EULAR response is the same as a DAS28 improvement of ≥1.2 – this is not accurate
2. The AG model assumes that if a biologic drug is withdrawn because of an AE within the first 6 months, then the patient receives rituximab – whereas NICE guidance states an alternative biologic should be considered

The AG model assumes that after their first biologic, all patients receive rituximab followed by tocilizumab – and NICE guidance does not explicitly state this.

A good or moderate EULAR response is not the same as a DAS28 improvement of ≥1.2

The Assessment Report states that:

> "…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." – Section 6.3, page 346.

The AG correctly state that NICE guidance has defined stopping rules where a biologic drug is stopped unless a DAS28 reduction of ≥ 1.2 points is achieved. However, the AG assumes that a moderate or good EULAR response is the same as a DAS28 improvement of at least 1.2 points, which is not a true representation of the relationship. The criteria for EULAR response is shown in Table 1.

Table 1: EULAR response categories

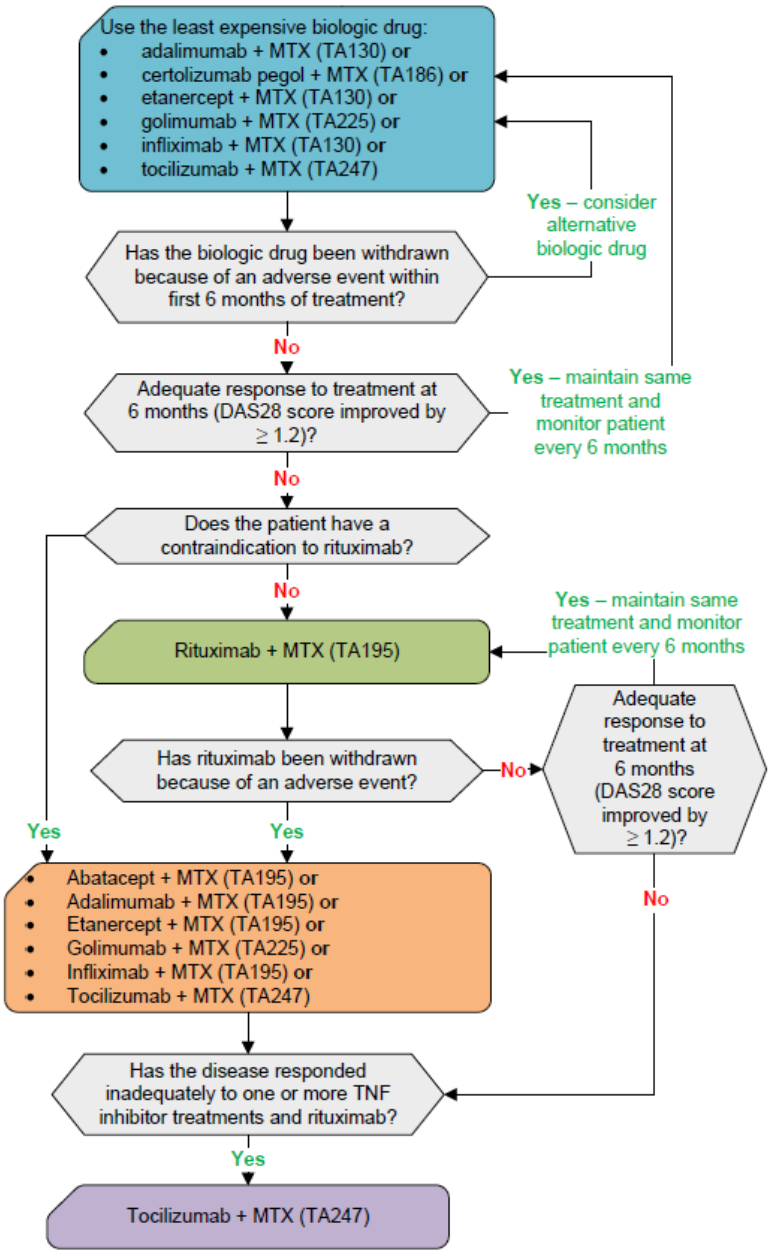
	Improvement in DAS28		
DAS28 at endpoint	**>1.2**	**>0.6 and ≤1.2**	**≤0.6**
≤ 3.2	good	moderate	none
>3.2 and ≤5.1	moderate	moderate	none
>5.1	moderate	none	none

Whilst it is true that a patient who has had a DAS28 improvement of greater than 1.2 will be classified as a good or moderate responder, it is not true that all patients classified as good or moderate responders have had a DAS28 improvement of at least 1.2. A patient can be classified as a moderate responder whilst achieving a DAS28 improvement between 0.6 and 1.2, if their DAS28 at endpoint is less than or equal to 5.1. By considering that all moderate and good responders according to EULAR criteria have had a DAS28 improvement of at least 1.2, it is likely that the AG overestimate the proportion of patients whose DAS28 score has improved by at least 1.2. The moderate responders who continue treatment in the AG model and would not in clinical practice due to insufficient DAS28 improvement, would receive fewer benefits from treatment for the same cost. This would increase the costs to a greater extent than the benefits, thus increasing the ICERs in the model. Indeed, in the scenario analysis conducted by BMS which considered that treatment continuation is dependent on good or moderate EULAR response (rather than DAS28 improvement of ≥1.2), the ICERs for bDMARDs compared to cDMARDs increased by an average of £2,600 (BMS submission, page 107).

Patients whose biologic drug has been withdrawn due to an adverse event at month 6 should consider an alternative biologic drug

The treatment algorithm for biologic drugs for RA at the time of the manufacturer submissions is reproduced in Figure 1 (this algorithm has since been updated (May 2013), but remains similar). This shows that if a biologic drug is withdrawn because of an adverse event by month 6, then an alternative biologic drug should be considered. In contrast, the AG model assumes the patient moves directly onto rituximab. Since rituximab is assumed by the AG to be similarly efficacious to the other biologics, but its annual cost is lower, the AG model potentially biases the results against the biologics, which exhibit more favourable tolerability profiles (although these are not considered in the AG model).

Figure 1: NICE biologics treatment pathway - February 2013



Patients should receive tocilizumab if their disease has responded inadequately to one or more TNF inhibitors and rituximab

The wording in the commissioning algorithm for biologic drugs for RA (Figure 1) asks whether the disease has responded "inadequately" to one or more TNF inhibitors and rituximab. Within the same algorithm, "adequate response to treatment at 6 months" is defined as DAS28 score improved by ≥1.2.

It would therefore seem reasonable to assume that patients would only receive tocilizumab if they had insufficient DAS28 improvement at month 6 – and that

patients who received rituximab long-term (indicating efficacy) should not receive tocilizumab.

Assessment Group model assumptions

There are a number of assumptions within the AG model that influence the results and may not be appropriate.

1. Assuming that HAQ improvement conditional on EULAR response is the same for all treatments could underestimate the effectiveness and hence cost-effectiveness of bDMARDs.
2. The HAQ trajectory analysis (assuming that the HAQ worsening on cDMARDs lessens over time) changes the AG model results by nearly 100%, but only changes the BMS model results by 30-40%
3. Assuming treatment effect on pain is directly proportional to effect on HAQ may not be appropriate and may double-count the HAQ benefit

Assuming HAQ improvement conditional on EULAR response is the same for all treatments may not be appropriate

The AG assume that HAQ change, conditional on EULAR response is equal irrespective of the treatment – whether a bDMARD or cDMARD.

It appears generally accepted that all bDMARDs are similarly efficacious. Many of the manufacturer models assumed that HAQ change was conditional on ACR response and was the same for all biologics. Given the similarity in the changes in HAQ scores from the trial data and BMS' MTC, it seems reasonable to assume that the HAQ change conditional on EULAR response is equal for all bDMARDs.

The AG analysed data for HAQ change conditional on EULAR response for cDMARDs from the ERAS dataset. This analysis showed that HAQ change conditional on EULAR response was greater for cDMARDS than the analysis for bDMARDs based on the BSRBR dataset. This is considered by the AG in a sensitivity analysis, but not in the base case. This suggests that the AG did not consider it reasonable that cDMARDs offered a greater HAQ improvement than bDMARDs, and so should raise questions about the comparability of the ERAS and BSRBR datasets.

If bDMARDS are more efficacious than cDMARDs, then assuming that HAQ change conditional on EULAR response is equal may overestimate the HAQ improvement for cDMARDs. EULAR response criteria classify patients based on their DAS28 improvement (and DAS28 score at endpoint). If we consider two patients who both have a starting DAS28 score of 4.4 before receiving different treatments, and one patient's DAS28 score improves 1.3, and the other's by 1.8, both patients will be classified as good responder. However, we would expect that the patient with the greater DAS28 improvement would also have a greater HAQ improvement.

Basing HAQ improvement on EULAR response would be fair if we believed that for every treatment, the mean HAQ improvement for moderate responders was the same, and the mean HAQ improvement for the good responders was the same. The

AG does not present evidence to justify this principle, which as a result may overestimate the HAQ benefit (and thus utility gain) for cDMARDs.

The HAQ trajectory analysis changes the AG model results by nearly 100%, but only changes the BMS model results by 30-40%

The AG found that changing the assumptions regarding HAQ trajectory made a substantial difference to the results, changing the ICERs for bDMARDs v cDMARDs by almost 100%. BMS therefore wanted to investigate the impact these assumptions would have on their cost-effectiveness model.

BMS has digitised and analysed the HAQ trajectories from the graphs published by Norton et al and in the 15 year study abstract. BMS recognise that our analysis is a simplification, and is limited as they cannot access the raw data. We have estimated the average HAQ increase between year 1 and year 10 from the ERAS dataset in the 15 year study, and found it to be 0.38. This equates to the time for the HAQ to increase by 0.125 being approximately 3.5 years (the BMS cost-effectiveness model considers time to HAQ increase of 0.125). BMS have assumed that the HAQ score remains constant after year 10, and does not increase further. We have considered a scenario in our economic model assuming that:

- HAQ increases on cDMARDs and palliative care if used within the first 10 years of cDMARD treatment
- HAQ increases at a rate of 0.125 every 3.5 years for cDMARDs, and 0.125 every 2 years for palliative care

This analysis increased the ICERs of bDMARDs compared to cDMARDs from BMS's base case, but not as dramatically as in the AG's analysis. The ICERs increase by around 30-40% - much less than the almost 100% increase seen by the AG analysis (Table 2).

Table 2: BMS sensitivity analysis - HAQ trajectory

	Discounted Costs	Discounted QALYs	ICER v cDMARDs
Certolizumab pegol	£[redacted]	9.71	£[redacted]
SC abatacept	£[redacted]	9.15	£[redacted]
Adalimumab	£[redacted]	9.02	£[redacted]
Golimumab	£[redacted]	9.02	£[redacted]
IV abatacept	£[redacted]	9.05	£[redacted]
Infliximab	£[redacted]	9.12	£[redacted]
Etanercept	£[redacted]	8.32	£[redacted]
Tocilizumab	£[redacted]	8.22	£[redacted]
DMARDs	£[redacted]	6.29	

Note: Highlighting denotes commercially in confidence information.

Although this analysis has its limitations, it demonstrates that the HAQ trajectory is not the only cause of the difference in results between the AG model and the BMS model. This may suggest that one or more of the other assumptions in AG model are interacting with the HAQ trajectory assumption. BMS are not convinced that many of these assumptions are appropriate, and an interaction between them would increase

the inaccuracy of the results from the AG model. As the costs and QALYs for the base case and scenario results are not visible within the Assessment Report, BMS are unable to try to comprehend the source of the differences in results both between the AG and BMS base case, and between the AG base case and HAQ trajectory sensitivity analysis.

Assuming treatment effect on pain is directly proportional to effect on HAQ may not be appropriate

The Assessment Report states that:

> "The fact that pain enters as a separate covariate in the Hernandez model is because HAQ and pain are not perfectly correlated. It is therefore important to include pain as an explanatory variable in estimating EQ-5D." – section 6.3.18, page 377

It is logical that including the effect of therapies on pain as we all as HAQ would improve the accuracy of the utilities within the economic model, and would help to distinguish between therapies which may have the same effect on HAQ, but a different effect on pain. It is plausible that this difference may exist – a therapy which relieves pain may make it less painful for a person to perform a given task, but not necessarily make it less difficult (as would be captured in the HAQ).

The AG, however, does not investigate the effect of therapies on pain, stating that the relationship between HAQ and pain can be incorporated in other ways without the requirement for additional complexity. It is surprising that the effect of therapies on pain is not investigated, given the anticipated improvement in the accuracy of the mapping. Instead, the AG estimates the expected pain score from the HAQ score, using an analysis from the NDB and ERAS datasets. The AG defends this with the statement that:

> "Whilst this assumes that all treatments affect pain proportionate to their effect on HAQ score this is also the assumption implicit in all models that exclude pain." – section 6.3.18 page 378.

This statement is not necessarily true. Excluding pain in the model assumes either that the effect on pain is not relevant in calculating utility, or that it is already incorporated into the HAQ score.

Assuming that a therapy's effect on pain is proportionate to its effect on HAQ potentially double counts the HAQ improvement. Different therapies may have different benefits in addition to HAQ improvement, such as a reduction in short-term discontinuation due to adverse events, and double counting the HAQ improvement potentially negates these other benefits.

The AG model does not distinguish between different tolerability profiles of the biologics

The AG assumes that the rate of serious infections observed is in 3.5% of patients on bDMARDs, and in 2.6% of patients on cDMARDs. This does not appear unreasonable for considering costs and disutilities, but it seems odd that the AG have not used the proportion of patients discontinuing treatments due to adverse events at 6 months from the trial data, which is reported for many of the RCTs.

The AG does not distinguish between the different tolerability profiles of the biologics. The adverse event profiles of therapies would be expected to be of interest to patients and clinicians, and would therefore be expected to be considered in modelling their effectiveness and cost-effectiveness.

Patient population

The patient populations for “moderate to severe” and “severe” are not defined in the scope, and the AG’s definitions are not the only plausible interpretation.

The patient populations in clinical trials used in the MTC to generate efficacy data are a mixture of moderate and severe patients. The AG has assumed the results from these trials apply equally to moderate patients and severe patients, which could lead to biased and inaccurate results.

The AG’s definitions of the patient populations are not the only plausible interpretations

The “moderate to severe” and “severe” populations were not defined in the NICE scope and were left to the interpretation of the submitting company. When a manufacturer (AbbVie) suggested stating the definitions explicitly in the scope, NICE responded “The scope is intended to provide guidance on the populations to be considered but their definition will be considered as part of the appraisal”. This response does not appear to discourage different interpretations of the populations.

The definition of “moderate to severe active RA” could be, and was by BMS, interpreted as including severe patients rather than excluding them. As such, BMS conducted their analysis within the eligible patient population – that is, patients with moderate or severe active RA entering clinical trials (typical inclusion criteria may be DAS28≥3.2, with no upper limit). BMS did not consider the subgroup of only severe patients, as efficacy data from clinical trials for all biologic drugs is not reported separately for this subgroup of patients.

The Assessment Report states that:

> “Although the NICE scope did not specify the definition of severe active RA and moderate to severe active RA, the following definition (based on expert clinical advice to the Assessment Group) has been adopted: severe active RA will be defined by a DAS28 score of ≥5.1, and moderate to severe active RA will be defined as a DAS28 score between 3.2 and 5.1.”

The AG therefore considers two separate, mutually exclusive subgroups for the population with “moderate to severe active RA” and “severe active RA.” The AG consider therefore that BMS’ submission did not consider either of the patient populations in the scope – we feel this criticism is not appropriate, given that the populations were not defined in the scope.

The AG model assumes that the efficacy results from a mixed population of moderate and severe patients can be applied equally to populations of only moderate, or only severe patients

The AG conducted an MTC on EULAR response, using the whole population from the MTC: the results are not broken down by “moderate to severe” or “severe” subgroups.

EULAR response is defined in terms of DAS28 improvement and DAS28 at endpoint. EULAR response must therefore also be influenced by DAS28 at startpoint (since DAS28 at endpoint = DAS28 at startpoint – DAS28 improvement).

The AG appears to have assumed, although they do not explicitly state so, that the probabilities of EULAR response from the mixed trial populations of moderate or severe patients apply equally to both the ‘moderate’ only patients, and the ‘severe’ only patients. This therefore assumes that the proportion of good responders in the moderate population is exactly the same as the proportion of good responders in the severe population, and the same for moderate and non-responders. This assumption is not necessarily true.

We would instead expect that there would be a higher proportion of good responders in the moderate population than in the severe population, since a severe patient would require a greater improvement in DAS28 to be categorised as a good responder. For example, consider a moderate patient with an initial DAS28 score of 3.2, and a severe patient with an initial DAS28 score of 5.1. The moderate patient will be categorised as a good responder if their DAS28 improvement is greater than 1.2, whereas the severe patient would have to have a DAS28 improvement of greater than 1.9.

Assuming that the same probabilities of EULAR response apply to the moderate and severe population is therefore likely to potentially overestimate the effectiveness in the severe population, and underestimate the effectiveness in the moderate population.

Conversely, it is plausible that a treatment may work better in the severe population than the moderate population (since severe patients have a greater potential for improvement), and so the effectiveness would be underestimated in the severe population and overestimated in the moderate population.

The effect of the AG’s assumption that the same probabilities of EULAR response apply to the moderate and severe population is unclear – but what is clear is that such an assumption could lead to biased and inaccurate results. This assumption is neither justified nor recognised as a potential limitation by the AG in their model.

Interpretation of ICERs for bDMARDs vs cDMARDs

The NICE scope stated that bDMARDs should be compared to each other and to cDMARDs. While the ICERs for bDMARDs v cDMARDs are useful for comparing the cost-effectiveness between the bDMARDS, they cannot be used to evaluate the cost-effectiveness of first-line bDMARDs compared to cDMARDs alone.

The scope stated that the bDMARDs should be compared to each other and compared to cDMARDs. The AG, like BMS, considered that incremental analysis between the bDMARDs may be misleading, because the ICERs for each bDMARD compared to cDMARD are relatively similar. As such, the ICERs for each bDMARD compared to cDMARD provide a useful platform for comparing the cost-effectiveness between the bDMARDs.

However, it is difficult to interpret the direct comparison between bDMARDs and cDMARDs, since the first-line bDMARDs are not actually compared to a cDMARD in the same positioning. What is actually compared is a sequence of three bDMARDs followed by cDMARDs alone, to cDMARDs alone – while this can be used to evaluate the cost-effectiveness of the bDMARDs treatment sequence compared to cDMARDs alone, it cannot be used to evaluate the cost-effectiveness of first-line bDMARDs compared to cDMARDs alone. Two alternative approaches could be considered which may be more meaningful, but do not match the treatment pathway:

1. We could consider a bDMARD followed immediately by cDMARDs (and not by rituximab or other bDMARDs) compared to cDMARDs alone. However, such a treatment sequence does not match NICE's recommended treatment pathway, and so does not address the questions posed in the scope.
2. We could consider a first bDMARD followed by second and third bDMARDs and then cDMARDs alone compared to a cDMARD followed by second and third bDMARDs and then cDMARDs alone. Again, this sequence does not match NICE's recommended treatment pathway, and rituximab would become a first-line biologic.

It is therefore difficult to make an appropriate decision based on the current scope.

ICERs for moderate and severe populations

We understand that the AG model uses the same efficacy and time on treatment data for the moderate and severe populations, but uses different baseline characteristics, such that the severe population have higher DAS28 scores, and correspondingly higher HAQ scores. The relationships between HAQ and hospitalisation costs and between HAQ and utility are non-linear, and so the cost and QALY gains for the moderate and severe populations are not the same. For patients experiencing the same improvement in HAQ score:

- The cost offset from reduction in hospitalisations for bDMARDs compared to cDMARDs is lower for the moderate population than for the severe

population, and so the incremental cost for bDMARDs versus cDMARDs is higher for the moderate population than the severe population.
- The QALY gain for bDMARDs compared to cDMARDs is lower for the moderate population than for the severe population, and so the incremental QALYs for bDMARDs versus cDMARDs are lower for the moderate population than the severe population.

Both of these factors contribute to the higher ICER for the moderate population than for the severe population.

Although we understand how the methodology employed in the AG model leads to higher ICERs for the moderate population than the severe population, we would have expected the results to indicate the opposite. We believe that the hospitalisation costs and QALY gains in the AG model are appropriate and reasonable and so we question whether the same efficacy data should be used for the moderate population and severe population.

The cost-effectiveness of treatment sequence after first-line biologics

If the cost-effectiveness of bDMARDs preceding an existing sequence of biologic treatments is to be compared to cDMARDs alone, the cost-effectiveness of the treatment sequence following first bDMARD should be considered. This subsequent treatment sequence has not been demonstrated to be cost-effective using the AG's model structure and assumption.

The scope stated that "The sequence of treatments used after the failure of the first biologic treatment should be included as part of the cost effectiveness modelling." Although it is outside the scope of this appraisal, if the cost-effectiveness of bDMARDs added to an existing sequence of biologic treatments is to be compared to cDMARDs alone, the cost-effectiveness of the sequence following first bDMARD should be considered.

The structure of, and assumptions in, the AG model differ greatly from those used to inform previous NICE recommendations for second- and third-line treatments. Although it is outside of the scope to consider the cost-effectiveness of second- and third-line treatments, and BMS recognise that better methods and data sources should be used where available, BMS suggest that the effect of the using the AG's assumptions and structure in later line models should be considered.

In particular, if the AG's assumptions for HAQ trajectory were used in the second- and third-line modelling, this would have increased the ICERs for rituximab and tocilizumab; potentially above NICE's commonly accepted thresholds. If these treatments were then not cost-effective, it would be difficult for a first-line biologic added onto this sequence to be cost-effective compared to cDMARDs.

Implication of not recommending bDMARDs

If first-line bDMARDs are not recommended, there will be implications for subsequent line treatments which have only been recommended (or appraised) in relation to bDMARDs at first-line. This would therefore change the scope of this appraisal.

Conclusion

BMS would like NICE to consider that the model developed by the AG has a number of flaws and limitations for evaluating the cost-effectiveness of bDMARDs in clinical practice. These come primarily from the differences between the model structure and the NICE recommended pathway. Furthermore, BMS are not convinced by the appropriateness of some of the assumptions used within the AG model, which are not supported by evidence. BMS request that these flaws/limitations and assumptions are revisited and assessed by NICE/the AG in order to produce a suitable tool for decision making in the ID537 assessment.

BMS also encourage NICE to consider the implication of the AG model assumptions within treatment sequences beyond first-line biologics.

Finally, we would like to reiterate our concern that this Assessment Group report indicates that no biologic therapies for the treatment of RA can be considered cost-effective as ‘The Assessment Group believes the ICERs for bDMARDs....is credibly greater than £50,000 per QALY when compared to a cDMARD alone strategy’; when previous NICE assessments, in the not too distant past, of these treatment options concluded that they are a cost-effective use of NHS resources and warranted positive HTA recommendations.