Schering-Plough Ltd
Shire Park, Welwyn Garden City
Hertfordshire, AL7 1TW
Tel: 01707 363636
Fax: 01707 363690

CONTAINED WITHIN:
ACD Response
Appendix 1: TAR Addendum Report Response
Appendix 2: HAQ multiplier: Simulated vs clinically observed
[redacted]

24th March 2010

Jeremy Powell
Technology Appraisal Project Manager
National Institute for Health and Clinical Excellence
MidCity Place, 71 High Holborn
London WC1V 6NA

Dear Mr. Powell:

RE: ADALIMUMAB, ETANERCEPT, INFLIXIMAB, RITUXIMAB AND ABATACEPT FOR THE TREATMENT OF RHEUMATOID ARTHRITIS AFTER THE FAILURE OF A TNF INHIBITOR (PART REVIEW OF NICE TECHNOLOGY APPRAISAL ("TA") GUIDANCE 36, TA 126 AND TA 141) – COMMENTS ON THE APPRAISAL CONSULTATION DOCUMENT ("ACD") AND ADDENDUM REPORT

Schering-Plough welcomes the opportunity to comment on the ACD which sets out the Appraisal Committee's ("the Committee") recommendations on adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis ("RA"). In addition to our ACD response, we have included a response to the Assessment Group's critique of the manufacturer submissions (the "Addendum Report") which was received by consultees after the first Technology Appraisal Committee Meeting on 4 February 2010. Prior to detailing our comments on the ACD, we wish to make a number of procedural points regarding this appraisal as follows:

Lack of transparency

Schering-Plough is particularly disappointed not to have had an opportunity to review and/or provide its responses to the Addendum Report prior to the first Committee meeting. Unlike members of the Committee who had the benefit of reading the Addendum Report prior to the meeting, Schering-Plough was unable to engage fairly in a balanced discussion with members of the Committee during that meeting. Had Schering-Plough received a copy of the Addendum Report, we believe that we could successfully have challenged the outputs of that report before the Committee. The failure to provide Schering-Plough and other consultees with a copy of the Addendum Report unfairly prejudices infliximab, particularly in the context of this appraisal where there has been a documented lack of transparency throughout.

Schering-Plough has not received a fully executable version of the model. We refer to our comments in our letter to you dated 12 January 2010, where we state that we have not been

able to validate the model given the near complete lack of explanation of the 2,000 lines of source code in the model. Until Schering-Plough has been given a fully executable version of the model, we are unable to scrutinise and validate it appropriately and therefore unable to engage effectively in consultation on the model or the ACD.

Failure to re-model and consider key evidence

Schering-Plough recognises the history of this appraisal since its inception within TA 130 over 5 years ago and aims to provide clinical and economic clarification of the evidence to assist the Committee in making its recommendations. It is highly regrettable that notwithstanding the recommendations of the Appeal Panel on the sequential use of adalimumab, etanercept and infliximab for RA, the Committee has yet to be presented with a comprehensive review of the available evidence, including relevant randomised controlled trial ("RCT") data and has failed to demand a re-modelling of the data. The Appeal Panel on sequential use said:

> *"The appeal panel considered that the topic should be re-scoped and that the Institute's normal procedures and methods, for a multi-technology assessment, should then follow. This should include invitations to consultees for submission of evidence,* ***re-modelling*** *if necessary and the development of new draft guidance for consultation." (Emphasis added.)*

Further, we are nonplussed by the failure to include key RCT data on tocilizumab, golimumab, and certolizumab pegol as these are comparators specifically referred to in the Final Scope of the appraisal. Failure to include such evidence, particularly as evidence incorporating the key trials was submitted by Schering-Plough and other consultees, is therefore outside the final scope of this appraisal and unfairly prejudices infliximab. We note the final protocol for this appraisal proposed a discretionary deadline for considering evidence relating to the above technologies, however, Schering-Plough considers that imposing such a deadline is itself unfair given that it restricts the agreed terms of the appraisal and that the manufacturers, who are ideally placed to inform NICE of likely marketing authorisation dates, were not consulted on this restriction. In any event, the deadline proposed was discretionary and given the relevance of the studies, the discretion should have been exercised in favour of including the studies in the modelling.

We urge the Committee to reconsider its preliminary recommendations in light of substantial clinical evidence and the alternative economic approaches submitted by Schering-Plough that reflect clinical practice in the UK, unlike the approach presented within the West Midlands Technology Assessment Report that does not reflect UK clinical practice.

We hope that following a review of our response, along with those of the other consultees, the Committee will re-model the data as recommended by the Appeal Panel above using a different Assessment Group that has not been involved in the review of tumour necrosis factor ("TNF") inhibitors for rheumatoid arthritis. Failing that, the Committee should require a re-evaluation of the approach and assumptions applied within the Birmingham Rheumatoid

Arthritis Model ("BRAM") and support a recommendation for the use of biological disease modifying anti-rheumatic drugs ("biologics"), including TNF inhibitors in the treatment of rheumatoid arthritis following an inadequate response on a first TNF inhibitor.

Schering-Plough has identified the following key points to inform the Committee that further assessment is required before proceeding beyond preliminary recommendations.

- **Substantial clinical evidence has not informed the appraisal**
  - Overly restrictive search criteria unfairly led to the exclusion of nearly 70% (113) of identified studies.
  - A detailed clinical write-up of GO-AFTER, the first RCT assessing the efficacy and safety of a TNFα inhibitor after an inadequate response to a first TNFα inhibitor has not been fairly or appropriately assessed by the Assessment Group and not given adequate consideration by the Committee.
  - Additional published or unpublished data (e.g., individual patient level data from the GO-AFTER trial) has not been requested from Schering-Plough by the Assessment Group or the Committee despite the clear need for such information. Analyses based upon GO-AFTER were misinterpreted within the Addendum Report and thus incorrectly questioned the validity of the submitted evidence.

- **The BRAM does not appropriately inform the decision problem. Alternative modelling approaches do exist which could better inform the Committee based on UK clinical practice**
  - Discrepancies exist between the simulated BRAM Health Assessment Questionnaire ("HAQ") multiplier and the HAQ multipliers observed from actual clinical trials. Schering-Plough has emphasised this in its response to the Assessment Report. The Assessment Group has made no attempt to validate this critical component, which is a primary determinant in the differential therapeutic effects between the biologics.
  - The BRAM relies on a health outcome with well-documented shortcomings, particularly the HAQ score. An alternative model submitted by Schering-Plough based on Disease Activity Score ("DAS"), which is more in line with UK clinical practice and NICE Guidelines, was unfairly dismissed purely because a small component of the model was informed by HAQ.
  - The BRAM does not include a treatment stopping rule based on a response criterion which is contradictory to previous appraisals (TA 130: 6 months stopping rule for TNFα inhibitors) and the current appraisal which found rituximab to be cost-effective based on the assumption that the product "is stopped if there is an inadequate response to treatment" (ACD, Section 4.3.22). The BRAM is highly sensitive to the stopping rule and Schering-Plough has a

legitimate expectation that the stopping rule would be included in this appraisal.

- **The appraisal does not take account of relevant safety information**
  - Preliminary recommendations have relied on the economic evaluation of rituximab and may not have fully taken into account the implications of solely recommending rituximab after an inadequate response on a TNFα inhibitor. Further to comments made in our response to the Assessment Report, Schering-Plough urges the Committee to consider the warning remarks issued by the Food and Drug Administration ("FDA") on the associated risk of treatment with rituximab and progressive multifocal leukoencephalopathy ("PML").

Detailed comments in relation to the ACD are presented below.

Exclusion of relevant clinical data: GO-AFTER Randomised Controlled Trial

GO-AFTER

"The Committee concluded that, although the studies suggest that a second TNF inhibitor is effective after the failure of the first, the absence of any rigorously controlled data meant that it could not quantify the relative effect of a second TNF inhibitor in comparison with either conventional DMARDs or alternative biological DMARDs" (ACD, Section 4.3.6, Page 35).

Indirect comparisons presented by the British Society for Rheumatology ("BSR"), two of the consultees as well as the West Midlands Assessment Group found no statistically significant difference in effect between the TNF inhibitors following an inadequate response to a first TNF inhibitor.

Similarly to the previous appraisal of TNF inhibitors for sequential therapy, where evidence for a biologic not being appraised was used to inform modelling by the Assessment Group (Schering-Plough TAR Response, Section 2.3.1), Schering-Plough considers the prospective, double-blind, placebo-controlled phase III trial that investigates the sequential use of TNF inhibitors in RA patients for Golimumab to be highly relevant to the decision problem. The Addendum Report unfairly disqualified our indirect comparison since it was presumed that we had included Golimumab trials not relevant to the TNF-experienced population (ie, methotrexate ("MTX")-naïve patients). By misrepresenting methods to the Committee, the Committee's conclusion that it was not appropriate to assume a class effect among the TNF inhibitors is misinformed (ACD, Section 4.3.5, Page 34). Schering-Plough is grateful now to have the opportunity to comment on the Addendum Report, following receipt of the document following the first Technology Appraisal Committee Meeting ("TAC") (note, however, our comments above on the lack of transparency).

In line with NICE's method guidance, which states that baseline utilities can be derived from other populations, Schering-Plough applied RCT data from GO-BEFORE and GO-FORWARD (Golimumab RCTs in MTX-naïve and MTX-experienced populations) solely to establish the baseline utility. The *relative* treatment effect was directly extracted from the GO-AFTER trial. Based on the robust and systematic indirect comparison, Schering-Plough concludes that the TNF inhibitors should be viewed as a class and therefore data from the GO-AFTER trial would be relevant to the decision problem. In any event, golimumab is listed as a relevant comparator in the Final Scope and evidence regarding golimumab should have been included in the appraisal.

Additional clinical evidence

Whilst RCT data is ideal, NICE's methods guide to technology appraisals states that other sources of evidence should be evaluated – particularly in light of available studies within a UK perspective. Large observational studies and registry data were dismissed but have large UK patient populations which can provide estimates of relative treatment effect. The Committee concluded that a recommendation for TNF inhibitors was not possible due to strict criteria relying ultimately on RCT data. Although relative efficacy and safety parameters from a large, observational study of adalimumab (n=899) by Bombardieri *et al* 2007 were presented within the Assessment Report, they appear to be dismissed in the final discussions.

Nearly 70% (113) of identified studies were excluded based on stringent and potentially arbitrary criteria (≥20 patients in an arm) (TAR, Section 5.1.2, Page 45). Biologics listed as treatment comparators within the Final Scope were not included within the search strategy (certolizumab pegol, tocilizumab, golimumab) and thus published findings (including RCTs) were not identified by the Assessment Group.

A cost-effectiveness analysis based solely on the BSR Biologics Registry ("BSRBR") data set was submitted to the Committee by the BSR for consideration. With a registry containing over 3,200 UK RA patients, Schering-Plough questions the reasons for dismissing this data based on perceived weaknesses in representing clinical practice. Given the absence of an alternative data set containing such large UK patient numbers and in light of an Appeal Panel Decision that suggested the appraisal "explain more fully its reasons for failing to recommend such treatment if there may be a reasonable possibility" (Appeal Panel Decision, TA130, Paragraph 141, Page 30), we would urge these data sets to be taken into account by the Committee. Our view is that there is sufficient data to form a recommendation for TNF inhibitors as a class.

[redacted]

The BRAM contains fundamental flaws and thus does not inform the decision problem

HAQ multiplier

"…Bearing in mind these considerations, the Committee accepted the use of a HAQ multiplier as a reasonable way to model changes in HAQ score" (ACD, Section 4.3.16, P41).

Lengthy discussions in the first TAC meeting on 4 February 2010 highlighted the issues regarding the HAQ multiplier: discrepancies between the simulated and clinical HAQ multipliers, weak base case data which was arbitrarily applied across multiple biologics, and general confusion on the applicability of the HAQ multiplier to clinical practice. It is therefore of some concern that the Committee has accepted this as the most appropriate method over the numerous alternatives presented by the consultees.

On 3 February 2010, Schering-Plough submitted additional evidence regarding the validation of the HAQ multiplier applied by the Assessment Group. This addendum was not included within the distributed Evaluation Report and may not have been taken into consideration by the Committee due to its late submission and thus is included again within *Appendix 2.*

As the HAQ multiplier is the primary determiner of differences in biologic treatment effects, the application of the simulated HAQ multiplier should be reflective of clinically observed outcomes. Given that *Appendix 2* graphically depicts the discrepancies between observed actual trials and the BRAM simulated HAQ multiplier, it is worrying that a major input of the BRAM is fundamentally flawed and this casts substantial doubt on the credibility of the resulting analysis.

Further, at the TAC meeting above, the Assessment Group responded to questions from the Committee about the failure to validate the simulated HAQ multipliers by stating that the Assessment Group had not been provided with the relevant data. Schering-Plough is willing to provide individual patient level data from appropriate trials if this would help the Assessment Group's analysis. However, the Assessment Group should have asked Schering-Plough for such data in accordance with NICE's usual procedures. To date, we have not been approached to provide further data, which is surprising given that other manufacturers in this appraisal have been asked to submit additional unpublished data. This inconsistent approach has unfairly prejudiced infliximab and led to perverse modelling.

HAQ vs DAS

"*…The Committee was mindful that all models presented had included EQ- 5D data derived from HAQ, and therefore no alternative was available...The Committee concluded that mapping HAQ to EQ5D had shortcomings, but in the absence of an alternative was an acceptable way to derive estimates of utility, and that the use of a non-linear function was not unreasonable" (ACD, Section 4.3.18, Page 42).*

The West Midlands Assessment Group has applied minor updates to the BRAM from the last appraisal rather than re-modelling and assessing whether an alternative approach may be better suited. The Committee notes that current clinical practice is shifting in line with NICE Guidelines and thus DAS may be a more appropriate health outcome measure than HAQ (ACD, Section 4.3.11, Page 37). Whilst the issues of HAQ within RA are documented extensively (ceiling effects, insensitivity at upper bounds for changes in quality of life and failing to fully capture treatment benefits [1,2]), the BRAM remains built around this inappropriate health outcome (ACD, Section 4.3.15, Page 40).

The Committee was informed that alternative approaches do not exist and that the manufacturers had all submitted models based on HAQ. However, this is misleading. Schering-Plough requests that the Committee considers our patient level model which only draws upon HAQ for a baseline response for conventional DMARDs, in line with NICE's recommended methodology. All relative treatment effects are driven by EULAR response (which is based on DAS). Schering-Plough recommends that the Committee requests a more robust alternative approach to explore appropriate modelling methods submitted by consultees, and which may be more in line with current NICE Guidance and therefore more suited to clinical practice.

Stopping Rule

"…The [BRAM] was not designed in a way which could incorporate stopping rules based on response criterion…The [BRAM] is most sensitive to changes in assumptions about…the number of people stopping treatment early" (ACD, 4.3.20, Page 44; 4.3.13, Page 39)

Four of the manufacturers submitted models that included a stopping rule based on treatment response. A sensitivity analysis conducted by the Assessment Group found that partial incorporation of a stopping rule based on response criterion within the BRAM lowered the ICERs by approximately £10,000 / QALY gained. The Committee found that the BRAM was not able fully to incorporate a response criterion. This coupled with a belief that the use of stopping rules does not reflect current clinical practice led to the Committee's conclusion that this component could not be the basis for decision making.

Schering-Plough believes that the Committee's view set out in section 4.3.20 of the ACD is perverse for the following reasons:

- Firstly, it is our understanding that the majority of Primary Care Trusts ("PCTs") will audit the use of TNF inhibitors and will require data on response. No evidence that we are aware of has been presented to substantiate the view that there is

[1] Aletaha D, Ward MM. Duration of rheumatoid arthritis influences the degree of functional improvement in clinical trials. Ann Rheum Dis 2006;65:227-233

[2] Scott DL, Pugner K, Kaarela K, et al. The links between joint damage and disability in rheumatoid arthritis. Rheumatology (Oxford). 2000 Feb;39(2):122-32.

widespread refusal to stop TNF inhibitor therapy where patients have not responded.

- Secondly, the Institute's own guidance as set out in TA130 requires that response to treatment is determined and non-responders are withdrawn from therapy. Modelling cost-effectiveness must therefore take this into account.
- Thirdly, it is unacceptable for the Committee to dismiss the relevant cost-effectiveness estimates on the basis that response criteria are only partially incorporated – clearly the appropriate course of action is to demand that the BRAM is amended to allow for this clinically important element of therapy to be fully incorporated, particularly since it appears that cost-effectiveness estimates are likely to be sensitive to this assumption.

Weaknesses within the structural aspects of the BRAM need to be addressed rather than dismissed, especially in light of cost-effectiveness arguments for the sole product recommended for sequential use, rituximab being conditional on the inclusion of a stopping rule.

In line with previous and ongoing appraisals, NICE recommendations have a notable impact upon prescribing patterns. Indeed, the current appraisal noted the effect of TA 130 guidance on treatment practice based on DAS response and tailored to the specified endpoints recommended in the final guidance. Schering-Plough urges the Committee to re-evaluate the strengths and challenge the shortcomings of the BRAM whilst working to apply solutions which are representative of the clinical practice that they envisage will comprise the best use of NHS resources. This may ultimately mean that the data needs to be re-modelled in line with the Appeal Panel's recommendation above, preferably by a different Assessment Group.

Further Comments on the ACD

Determining response to treatment in the Schering-Plough model

Section 4.2.12 of the ACD summarises how response to treatment is determined in Schering-Plough's model. The summary provided is somewhat misleading. The first step in the two-step process is more accurately defined as follows: baseline EULAR response data from the BSRBR (from TNF inhibitor experienced DMARD receiving patients) was converted to baseline ACR response using an algorithm derived from the GO-AFTER trial, results from the MTC on the ACR scale were then applied to generate ACR responses for each treatment, these were then converted back to EULAR response rates.

Sub-group analysis: contraindication or intolerance to rituximab

Section 4.3.25 of the ACD states that the Committee considered that it had not been presented with any clinical evidence regarding the use of TNF inhibitors or abatacept in patients for whom rituximab failed or in whom rituximab was contraindicated on not tolerated. However,

as set out elsewhere in this response, there is a large body of evidence demonstrating the effectiveness of TNF inhibitors used sequentially. It is not clear why the Committee believes that it requires further specific evidence in relation to patients with for example a contraindication to rituximab, particularly given that it acknowledges that ICERs presented for TNF inhibitors compared with conventional DMARDs are a reasonable proxy for this sub-group analysis. Schering-Plough requests that the Committee reconsiders its assessment of this potential patient group and is more explicit about its rationale for not giving adequate consideration to the potential cost-effectiveness of TNF inhibitors used in these patients.

Subgroups based on the presence of auto-antibodies

The Committee concluded that there was insufficient evidence to make differential recommendations for subgroups based on auto-antibody status. Schering-Plough believes that the Committee has not given adequate consideration to the reduced or absent response of seronegative patients to rituximab. The Committees view appears to have been determined on the basis that there were no statistical differences observed in trials designed with seropositivity as a selection criterion and that were inadequately powered to test a hypothesis regarding this issue. Independent data showing reduced or no responses in patients who are seronegative are available.[3] Schering-Plough believes that the Committee has failed to give adequate consideration to this patient population.

Relevant safety information not included

As raised and discussed in the previous TAC, concerns exist over the safety profile for rituximab. Schering-Plough urges the Committee fully to consider the risk of PML as detailed within our Assessment Report response (Schering-Plough TAR response, Section 2.1.1, Page 4).

Vial Optimisation

Vial optimisation with infliximab in RA has implications on the cost-effectiveness of the technology (see comments in our letter to you dated 12 January 2010). Following the Appeal to TA130, the Committee was instructed to consider an appropriate range of doses for infliximab and to take account of vial wastage. Further, NICE's response to Schering-Plough's comments on the Draft Scope for this appraisal explicitly state that a range of doses will be taken into account. NICE stated: "*All included technologies will be appraised as per their respective licensed indications, which will include the alternative dosing schedules for infliximab. No changes made to the scope.*"[4] The failure to take account of vial optimisation is unfair and outside the Final Scope given the comments above.

In its addendum to the Assessment Report, the Assessment Group notes that "In any case all [*sic*] any savings from vial sharing are dwarfed by dose escalation. In the cited systematic review 44% of patients treated with infliximab had the drug dose increased." This is apparently a justification for not accounting for vial optimisation in the economic evaluation.

[3] Personal Communication: Dr T Hammond, Chair of Kent and Medway Rheumatology Steering Group
[4] See http://www.nice.org.uk/nicemedia/pdf/RheumatoidArthritisComments.pdf

In relation to this, we would like to raise two fundamental issues – firstly, the evidence presented for dose escalation and the extent to which it "dwarfs" vial sharing is sourced from outside the UK – the majority of studies were from the USA and all studies were published between 1998 and 2002; secondly TA130 recommends against dose escalation and an economic evaluation to determine the effective use of NHS resources ought to reflect this. Further, a recent update to the research conducted by Schering-Plough, with a higher response rate (57%) compared to that referred to by the Assessment Group, confirms similar findings to those reported in our original submission and research – i.e. that around two thirds of patients receive infliximab which has been prepared using vial optimisation.

Conclusion

The preliminary recommendations by the Committee do not take into account all of the available evidence to inform this appraisal. Manufacturers were not given the opportunity to comment on the Addendum Report which misinterprets much of the evidence and misinforms the Committee. Sufficient clinical evidence exists for the Committee to form a recommendation for the sequential use of TNF inhibitors for the treatment of RA. Given the considerable attention that has been paid to identifying all potential sources of evidence for sequential therapy during the course of TA130 and the Appeal Hearing of 29 September 2008 regarding sequential use of TNF inhibitors, it is unfair and perverse to assess biologic DMARDs without the inclusion of all of the available evidence for potentially relevant comparators, particularly given the fundamental flaws and lack of transparency over the BRAM model.

Based on the concerns raised above, Schering-Plough questions the validity of the conclusions reached by the Committee within the ACD and believes that substantial adjustments, if not a complete re-modelling, are needed by this Assessment Group or a different Assessment Group for the appraisal to reflect clinical evidence and fully inform the Committee.

We are grateful for the opportunity to comment on the ACD and the Addendum Report and look forward to continued dialogue with NICE regarding the issues raised in this response. Please do not hesitate to request any additional data from us which may be of use during this appraisal.

Kind regards,

[redacted]
[redacted], Market Access
Schering-Plough

RE: ADALIMUMAB, ETANERCEPT, INFLIXIMAB, RITUXIMAB AND ABATACEPT FOR THE TREATMENT OF RHEUMATOID ARTHRITIS AFTER THE FAILURE OF A TNF INHIBITOR (PART REVIEW OF NICE TECHNOLOGY APPRAISAL ("TA") GUIDANCE 36, TA 126 AND TA 141) – TECHNOLOGY APPRAISAL REPORT ("TAR") ADDENDUM

We welcome the opportunity to comment on the Addendum Report which was received prior to the first technology appraisal committee. Schering-Plough urges the Committee to consider the below misinterpretations of submitted evidence that we have highlighted below and which we believe may have misinformed the preliminary recommendations within the ACD.

MTCs in patient populations outside the scope
Addendum: Section 3, P22-23
"Due to the broad inclusion criteria beyond the scope of the appraisal, substantial clinical and statistical heterogeneity exists between the RCTs included in the MTCs. The basic requirement for IC/MTCs regarding the exchangeability of relative treatment effects between the included studies could not be assumed and thus the validity of the results was questionable. The violation of the basic requirement was particularly prominent in the MTCs conducted by Schering-Plough and Abbott...The substantial heterogeneity among studies in the MTC and the discrepancy between the results from these analyses and those actually observed in RCTs raise serious concern with regard to the validity of the MTC as well as the validity of the economic evaluation that utilised data from them"

Schering-Plough's model did include prior treatment duration as a covariable to adjust for differences between trials. Direct and indirect estimates of treatment effects were also compared and were found to be in good agreement. The Assessment Group's analysis appears to have compared the predicted and observed absolute probabilities of ACR response for individual trial arms rather than comparing the predicted and observed relative treatment effects. This is analogous to basing an indirect comparison on individual arms rather than the within trial relative treatment effects.

MTCs in patient populations outside the scope
Addendum: Section 3, Page 23
"Despite the broad inclusion criteria for the MTCs, clinical and methodological similarity/difference between included studies was only briefly described or not mentioned at all. Statistical heterogeneity between included studies were either not assessed or (where assessed) only dealt with by using random effects model without further exploration of potential source of heterogeneity. All the MTCs included a head-to-head trial (ATTEST, comparing infliximab to abatacept) but did not examine the direct evidence separately from indirect evidence. Consistency between direct and indirect evidence was not examined."

As a point of clarification, the MTC analysis did include a meta-regression with mean duration of prior therapy as a covariable in an attempt to account for heterogeneity and did compare the results of direct and indirect comparisons. We also note that there are no published trials of the TNF inhibitors infliximab, adalimumab and etanercept in a population of patients who have failed a previous TNF inhibitor. We did assume that the relative treatment effects seen on the

odds ratio scale would have been the same in TNF inhibitor-naive patients and TNF inhibitor-experienced patients. There is supporting empirical evidence for this assumption. However, the relative treatment effects were applied to a lower baseline response leading to a reduced absolute response. In addition, the model accounted for longer prior disease duration in these patients further reducing both the relative and absolute response.

Characteristics of starting population

Addendum: Section 4.3, Page 36

"The characteristics of starting population were based on GO-AFTER (a golimumab trial in patients who had inadequate response to TNF inhibitors): mean age 54, female 79%, baseline HAQ 1.61. The starting population was younger and had much lower baseline HAQ score compared to corresponding patients in BSRBR. Baseline utility (EQ-5D and SF-6D) was imputed from baseline HAQ using simple linear regression (lower HAQ corresponding to higher utility). The consequence is that the estimated baseline utility may have been higher than it should be."

The GO-AFTER study was used to estimate baseline utility as Schering-Plough had access to the necessary patient level data and the study was in an appropriate population. The estimated response to treatment was *incremental* to this initial utility and as such it is unlikely that the estimated ICER would be highly sensitive to the estimate of baseline utility. If requested, Schering-Plough would have carried out a confirmatory sensitivity analysis.

Estimates of clinical effectiveness – short term

Addendum: Section 4.3, Page 37

"EULAR response for corresponding patients who switched to a second TNF inhibitor (rather than conventional DMARDs) was available from the same analysis but this data was not used in the model. Instead estimates of effectiveness for TNF inhibitors were taken from the MTC and thus the data for comparative effectiveness were obtained from different sources that may not be comparable."

The EULAR analysis of incremental treatment effects did not distinguish between individual biologics and was based solely on observational data. The submitted analysis attempted to make use of RCT data in estimate incremental treatment effects. This corresponds to standard practice whereby estimates of baseline severity or event rates are often obtained from observational studies and estimates of incremental treatment effects are obtained from synthesis of RCT data.

As a point of information: a cost-effectiveness analysis based solely on the BSRBR data set was submitted to the committee by the BSR for consideration.

Estimates of clinical effectiveness - long-term

Addendum: Section 4.3, Page 37

For patients receiving biologics, the base case analysis assumed zero utility progression. A sensitivity analysis was carried out assuming utility progression was equal to that observed in the BSRBR (by EULAR response), which suggests that utility worsens for EULAR good responders, is close to zero for moderate responders and improves marginally for none responders. This seems counter-intuitive."

This was a sensitivity analysis based on empirical data from the BSRBR. The apparent counter-intuitive results may reflect regression to the mean over time.

Discontinuation rule and treatment duration

Addendum: Section 4.3, Page 38

"Treatment withdrawals were assumed to be the same for rituximab and abatacept. This assumption may over-estimate the proportions of people who continue with these therapies although data are limited."

The data on drop-out rate are indeed limited and in the absence of evidence that drop-out rates are either better or worse for anti-TNFs we set these to be equal across the biologics. If requested, Schering-Plough would have conducted further sensitivity analyses.

Appendix 2

Schering-Plough Ltd
Shire Park, Welwyn Garden City
Hertfordshire, AL7 1TW

Tel: 01707 363636
Fax: 01707 363690

3rd February 2010

Jeremy Powell
Technology Appraisal Project Manager
National Institute for Health and Clinical Excellence
MidCity Place, 71 High Holborn
London
WC1V 6NA

Dear Jeremy,

ADDENDUM TO SCHERING-PLOUGH'S RESPONSE TO THE WEST MIDLANDS HEALTH TECHNOLOGY ASSESSMENT COLLABORATION (WMHTAC) ASSESSMENT REPORT ON ADALIMUMAB, ETANERCEPT, INFLIXIMAB, RITUXIMAB AND ABATACEPT FOR THE TREATMENT OF RHEUMATOID ARTHRITIS AFTER THE FAILURE OF A TNF INHIBITOR

Further to our letter of 12th January, setting out Schering-Plough's formal response to the WMHTAC Assessment Report, I am writing to clarify certain aspects of the information set out in that letter.

In Section 2.3.3 of our letter of January 12th, we included an assessment of the BRAM – specifically regarding the use of the 'HAQ multiplier'. The HAQ multiplier has been described on a number of occasions by the WMHTAC and to our knowledge it has been a core feature of the BRAM since its inception.

Schering-Plough has now conducted further analyses with regard to the HAQ multiplier. As a result we have determined that figures presented in our letter of January 12th are inaccurate. Corrected figures replacing those in our original letter are presented below. However, the key concern raised by Schering-Plough remains – i.e. since the BRAM relies on the HAQ multiplier as a key determinant of clinical- and therefore cost-effectiveness, the model lacks face-validity.

The observed HAQ Multipliers from ASPIRE, ATTRACT and GO-AFTER are shown on the figures below, alongside HAQ Multipliers from the BRAM.

Figure 1

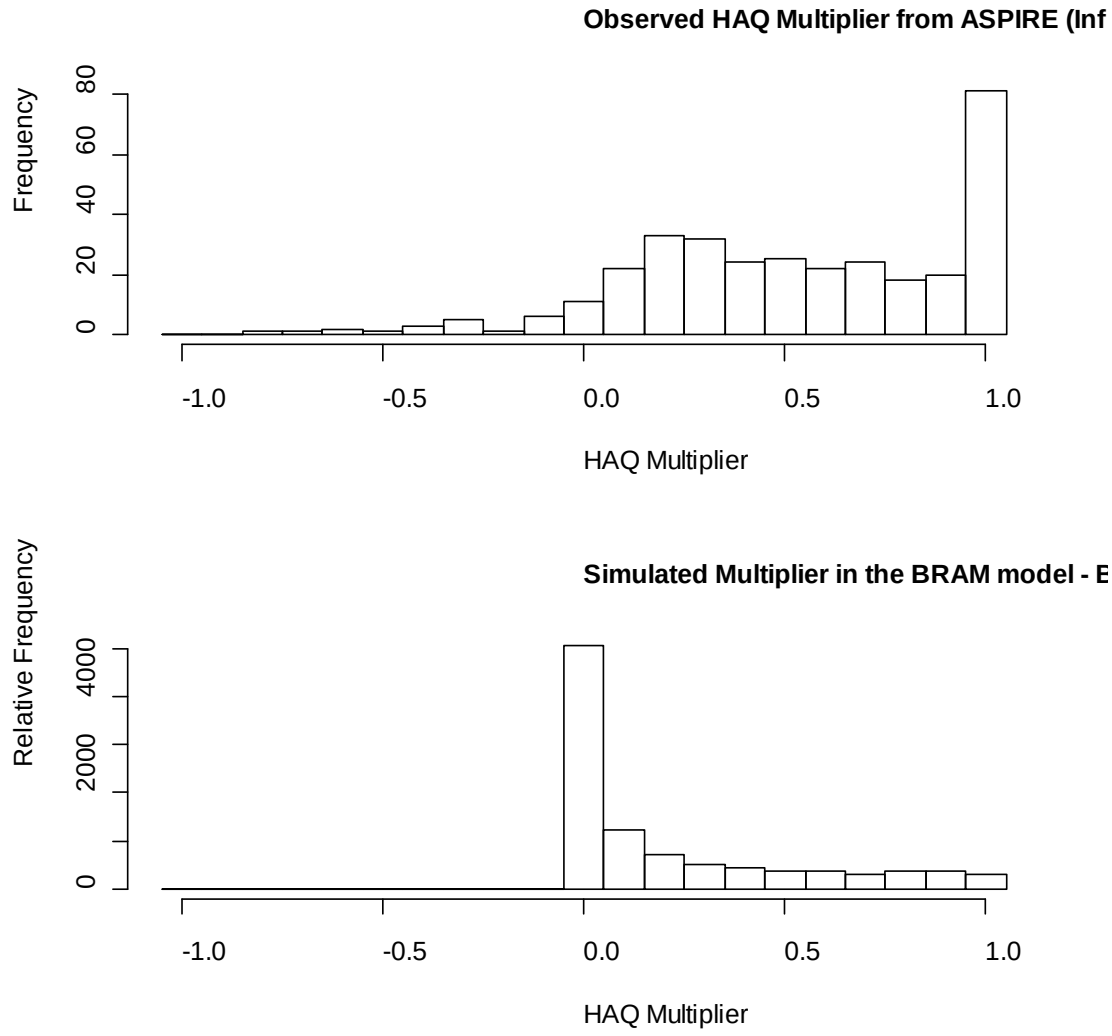

Observed HAQ Multiplier from ASPIRE (Inf
Frequency
0
20
40
60
80
-1.0
-0.5
0.0
0.5
1.0
HAQ Multiplier
Simulated Multiplier in the BRAM model - E
Relative Frequency
0
2000
4000
-1.0
-0.5
0.0
0.5
1.0
HAQ Multiplier

Figure 2

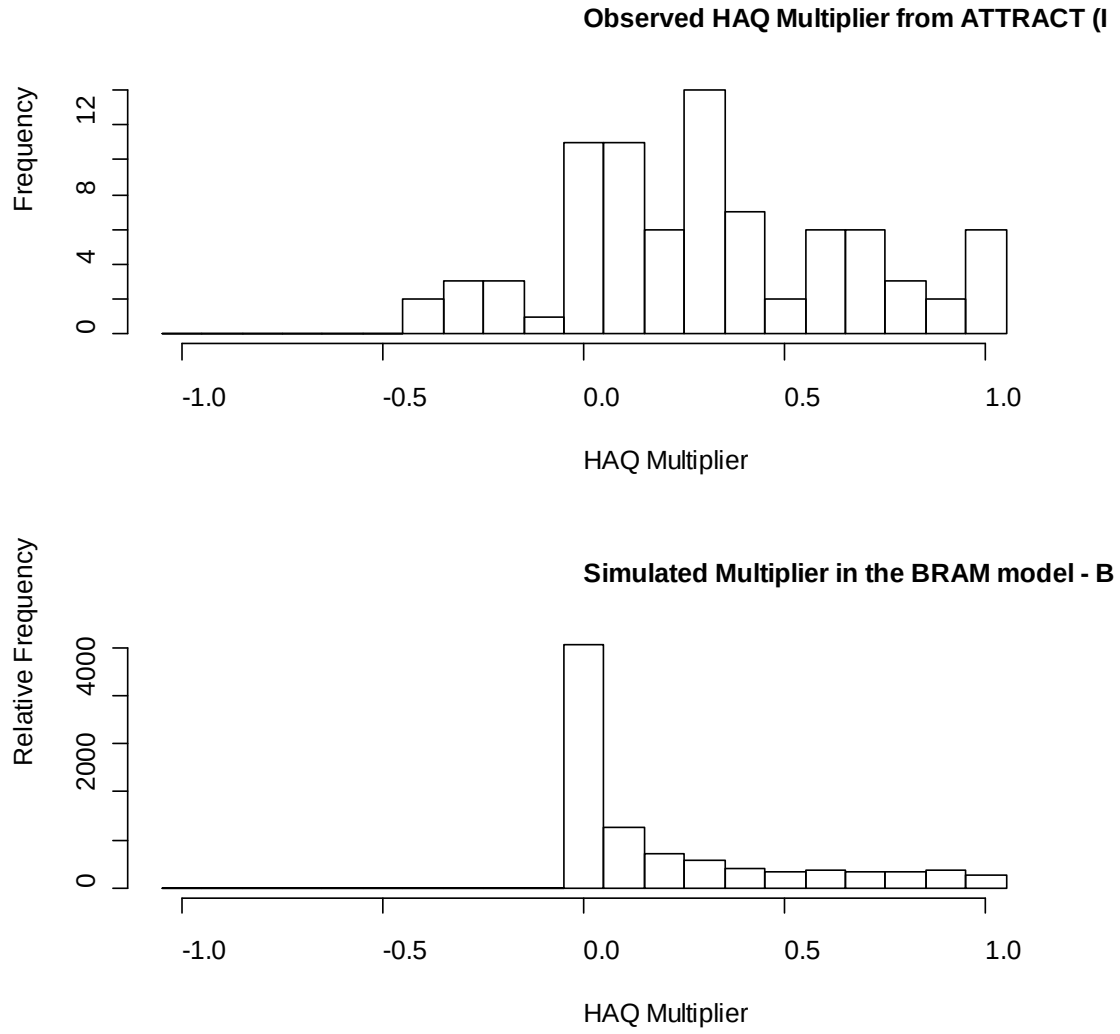
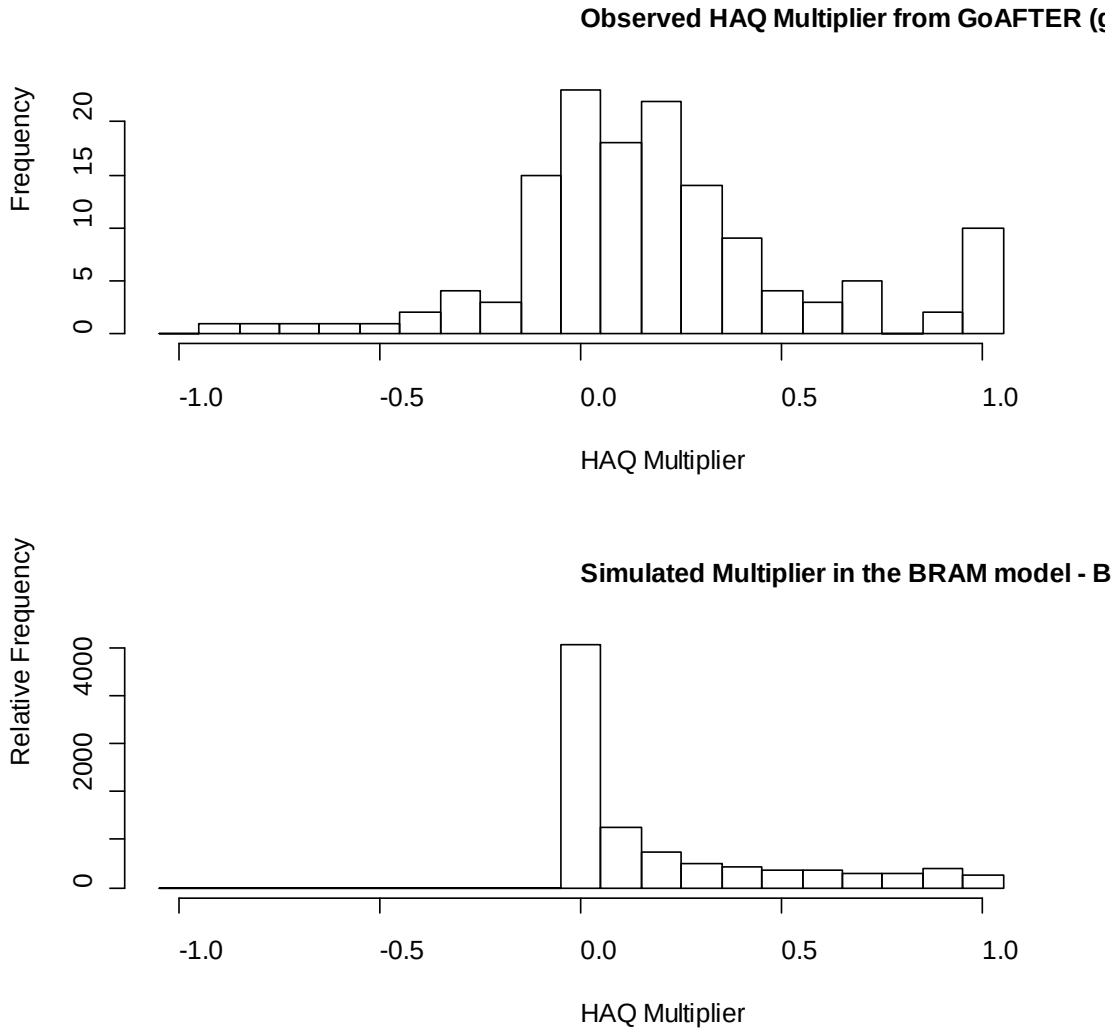

Observed HAQ Multiplier from ATTRACT (I
Frequency
0
4
8
12
-1.0
-0.5
0.0
0.5
1.0
HAQ Multiplier
Simulated Multiplier in the BRAM model - B
Relative Frequency
0
2000
4000
-1.0
-0.5
0.0
0.5
1.0
HAQ Multiplier

Figure 3


Observed HAQ Multiplier from GoAFTER (
Frequency
0
5
10
15
20
-1.0
-0.5
0.0
0.5
1.0
HAQ Multiplier
Simulated Multiplier in the BRAM model - B
Relative Frequency
0
2000
4000
-1.0
-0.5
0.0
0.5
1.0
HAQ Multiplier

The simulated HAQ multiplier used in the BRAM model does not appear to provide a good approximation of the results observed in actual trials. The model does not appear to have face validity with respect to its most critical component, the HAQ multiplier.

Whereas the HAQ multiplier used in the BRAM generates a high proportion of values close to zero, which represent very small or zero reductions on HAQ, the observed data from three clinical trials demonstrate a very different distribution of the HAQ multiplier with a spread of results ranging from negative values (indicating an increase in HAQ) up to 1 (indicating a reduction in HAQ to zero).

Schering-Plough considers this issue to be of potentially critical importance to the validity of the BRAM model and we request that these additional analyses are brought to the Assessment Group and Appraisal Committee's attention.

Sincerely,
Kind regards

[redacted]

[redacted], Market Access & Communications
Schering-Plough