AbbVie's Comments on the report by the DSU evaluating HAQ progression in people with RA treated with non-biologic therapies and on the addendum to the Assessment Group's report assessing how the results from the DSU report on HAQ progression affect the cost per QALY gained ratios

NICE MTA review of:

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)

AbbVie welcomes the opportunity to comment on the report by the Decision Support Unit (DSU) evaluating any additional evidence relating to the rate of HAQ progression of RA patients not treated with biologic therapies as well as the addendum to the Assessment Group's report assessing how the results from the DSU report on HAQ progression affect the cost per QALY gained ratios.

AbbVie understands that this consultation is limited to these two documents only, but would like to reiterate that all the issues raised by AbbVie during consultation of the original Assessment Group (AG) Report in August/September 2013 are still valid and should be considered by the Appraisal Committee in conjunction with these comments. AbbVie welcomes the fact that some changes have been incorporated in to the AG model as a result of comments from Consultees, but considers that there are additional points raised in the response to the original AG document submitted by AbbVie 23 September 2013 that have not been addressed yet.

AbbVie's response to this consultation is split into two parts: Section 1 focuses on the DSU report and Section 2 focuses on the addendum to the AG report.

1 AbbVie comments on DSU report

The DSU report is a comprehensive analysis investigating any additional data sources to inform the estimates for HAQ progression in RA patients treated with non-biologic agents. However, there are a number of considerable issues with the data sources that mean that the evidence used to inform the latent class growth model to estimate HAQ progression in the AG report is not robust, and more importantly, neither is it reflective of disease progression in the relevant RA population.

1.1 Comments on Section 2 of DSU report

Of the identified nine studies in Norton systematic review of studies that reported on HAQ change over time, the first major issue is that patients had either RA or undifferentiated polyarthritis. It is therefore not clear what type of RA patients were included i.e. in terms of severity or classification; if they had RA at all; or whether the proportion of undifferentiated polyarthritis affects HAQ progression rates. Regardless of this limitation, interestingly the HAQ progression rates reported across these studies for patients on cDMARDs are either less than the 0.045 rate assumed in the manufacturer models, 0.045 (e.g. Hawley and Wolfe), or in some cases greater – for example Leigh et al showed that disease duration matters in their study; RA patients > 20 years disease experienced faster deterioration than those patients with < 20 years disease. This study is in complete contrast to the data in the AG report that assumes that HAQ deterioration slows over time. AbbVie considers this is probably because the ERAS dataset utilised by the AG does not collect data for a long enough period of time, utilises far too broad an RA population, and has an extraordinarily high attrition rate that is difficult to meaningfully adjust for (see Section 1.2 and 1.3).

All the studies in the systematic review suggest that the rate of HAQ progression is dependent on certain disease characteristics. The DSU report is in agreement with this in that the latent class model is devised using different factors at baseline to predict HAQ trajectory. However, given the limitation with the data sources (discussed more in Section 1.2 below), AbbVie considers that not all the relevant factors have been taken into account and that the inception cohorts in their entirety are not predictive of RA patients with active moderate to severe disease who have failed 2 DMARDs.

1.2 Comments on Section 3 and 4 of DSU report

In this additional work undertaken by the DSU, several datasets were identified as potential sources for estimating cDMARD HAQ progression as alternative sources to the ERAS data that informed the latent class growth model (LCGM). With the exception of the US NDB, the four data sources identified are inception cohorts – BARFOT, LEC, ERAN and ERAS. AbbVie considers that it is the use of inception cohort data to inform the latent class growth models which are then used to predict HAQ progression for RA patients with active disease despite treatment with 2 or more cDMARDs that is the issue rather than the adoption of a LCGM. There are a number of reasons why an inception cohort is likely to underestimate the true HAQ progression for the population of interest – active moderate to severe RA despite cDMARD treatment:

1.2.1 Inclusion of early and self-limiting RA patients

Firstly, 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 would underestimate HAQ progression for the select group of patients who should be eligible for biologic therapy. AbbVie notes that the AG validated the HAQ progression estimation based on ERAS data with different datasets (based on NOAR and the ERAN – page 21 of DSU report), but similarly to the limitations of ERAS, these two cohorts include patients with mild RA, who would be ineligible for biologics therapy. Therefore, the validation of the ERAS dataset by ERAN and NOAR is almost irrelevant in that the validation has been corroborated in the same 'incorrect' RA population – early RA disease pre-cDMARD treatment, including those with mild or self-limiting disease.

Furthermore, page 26 of the DSU report indicates that only 53% of patients and 70% of patients in the ERAN and ERAS cohorts fulfilled the American College of Rheumatology's (ACR) criteria for RA, respectively. This suggests that a large proportion of patients may have undifferentiated had inflammatory polyarthritis; a condition with a different prognosis to RA which will affect the overall HAQ progression, especially given that ACR criteria was one of the characteristics used to predict latent class membership in the model.

In support of the assumption that HAQ progression differs in early and established disease, AbbVie conducted a latent class growth analysis similar to Norton *et al* .using data from the CORRONA RA registry to examine the differences in 5-year HAQ progression between patients with early and established RA. Included patients had no prior exposure to biologics at baseline and at least two years post index date (censored after initiating biologics), were in moderate to high disease severity, and used at least two cDMARDs one of which was methotrexate. The analysis showed considerable variations in HAQ progression trajectories between early and established RA patients. Among established RA patients (RA disease duration >2 years), the best fitting group-based model was with six trajectories compared to early RA patients model with only five trajectories, compared to the Norton model which indicated that four classes produced the best fit. This analysis demonstrates the importance of studying HAQ progression separately in early and established RA patients, and therefore data from ERAS – an inception cohort is not predictive of HAQ progression for patients with established active disease who have failed cDMARD treatment.

1.2.2 Majority of population informing HAQ progression rate are cDMARD responders

Secondly and crucially, during the entire observation period of the cohorts e.g. observations were available up to 15 years for ERAS, only 26% (ERAS) and 28% (ERAN) of patients failed 2 DMARDs <u>during</u> the observation period (Table 4 in DSU report). This means that the remaining 72-74% of patients had not failed 2 DMARDs and were still responding to their DMARD treatment, and it is these data that are driving HAQ progression for patients with active disease despite cDMARD failure in the AG model. It is an untenable assumption that

over 70% of the dataset who are still responding to DMARDs can be utilised to inform HAQ trajectories in a cDMARD failure population. It is this assumption that cDMARD failure has no impact on HAQ trajectory that leads to the underestimation of HAQ progression for cDMARDs in the model, because DMARD failure status is not a predictor for latent class membership. Particularly when it has been well documented in terms of treatment response, that previous number of DMARDs failed impacts the probability of response for a treatment - most probably because these patients are more difficult to treat and have prognostically more severe, and therefore more progressive disease.

To further test if failing cDMARDs has any impact on HAQ progression, as a part of the analysis of the CORRONA database, AbbVie compared HAQ trajectories from two groups of RA patients much more relevant to this appraisal than the ERAS dataset: those that had failed two cDMARDs vs. those that had failed three or more cDMARDs. The analysis showed that patients who had failed three or more cDMARDs were in higher HAQ-based trajectory groups compared to patients who failed only two cDMARDs. These findings combined with the sub-group analysis in Section 4.5 of the DSU report indicate that there is a definite trend regarding the number of DMARDs failed and the rate of HAQ progression.

To take the DSU's assessment of the suitability of the cohorts to the decision problem one step further, only 10% and 11% of the ERAS and ERAN cohorts fulfil the current NICE criteria for a biologic (Table 4 in DSU report). This means that the vast majority of patients in the ERAS cohort used to estimate HAQ progression in the AG report are not representative of the patient population eligible for a biologic under current NICE guidance. Therefore, regardless of how well the LCGM fits, it is the use of the ERAS data set in its entirety to generate the four latent classes that is inappropriate.

1.2.3 Impact of extremely high attrition rate in cohorts

Thirdly, and not unexpectedly there is an extraordinarily high drop-out/attrition rate for these cohorts over the observation period. Tables 8 to 11 of the DSU report highlight the sample sizes at the different time points in the cohorts. For the ERAN cohort, after 11 years of observation only 0.53% of the original sample remains to estimate HAQ; in fact from year 7 onwards less than 10% of the original sample is available for analysis due to attrition. For the ERAS cohort – the cohort used to generate the latent class growth model - at t=15 years only 2.45% of the original sample remain, and indeed from around 10 years since the start of the cohort less than 20% of the original sample remain to contribute observations. Both the LEC and BARFOT cohorts state that remission is a component of leaving the study, and as such these cohorts are potentially not as relevant as other data sources to estimate HAQ progression. However, the fact that a proportion of these patients are in remission from cDMARD use only confirms that these inception cohorts are not the relevant RA population to be estimating HAQ progression for cDMARD failure RA patients with active disease. Finally, the NDB is a massive database with a large number of observations e.g. starting sample size was n=19,462. However, only 3 years into the observation period and the sample size was only 25% of original sample (n=5,011). By 10 and 15 years, respectively, only 5% and 0.82% of the original sample remain to contribute to the HAQ data.

The DSU notes that whilst there is some censorship (particularly for the NDB – around 7%), the main reason for drop-out is attrition. This is very important for the cohorts collecting data during the era of biologics e.g. the NDB, ERAN, BARFOT etc., as it is highly likely that patients either dropped out of the cDMARD arm due to remission (therefore they are not the appropriate population anyway) or because they had active, progressive disease and were moved on to a biologic. Therefore, any patients remaining in any one of the above mentioned cohorts represent a biased group of cDMARD responders that are self-selected to remain in the cohort because they are neither in remission nor have they failed treatment, and are certainly not representative of RA patients who would be treated with biologics in the

UK. This will bias HAQ observations and is likely to be responsible for the flattening of the progression rate over time, as it appears that only a certain type of DMARD responder remains in the cohort beyond around 5 years. AbbVie contends that the non-linear progression rates observed in these data sets are as a result of selection bias selecting DMARD responders who remain in the cohort and not being able to adequately account for the high attrition rate.

AbbVie's conclusions are supported by data from the CORRONA database which show that up to 32.0% of RA patients switched from cDMARDs to a biologic between years 3 to 5 from t=0 due to active progressive disease.

1.2.4 Handling of missing data

The DSU acknowledges that there is an extraordinarily large drop-out in the data analyses and attempts to account for this using not missing at random (NMAR) techniques (Section 4.5 of the DSU report). The DSU focused on the ERAS cohort because it has a larger number of observations than the other inception cohorts and the data were also collected before the use of biologic drugs was common, meaning that fewer of the patients would have been removed from the sample because they were on biologics. Furthermore, it is the dataset that informs the latent class model.

Several of the NMAR techniques resulted in non-convergence in the model, and were therefore not useful – this is not surprising given the sheer volume of dropout discussed in 1.2.3 above. All the NMAR methodologies employed were linked to the latent class structure, which means that the baseline characteristics used to predict membership to a given latent class will ultimately also affect the NMAR imputations too, which is not ideal given AbbVie's concerns with the populations used to inform class membership mentioned so far. Results from the Roy-Muthen method show that unsurprisingly, those patients that are likely to drop out early in the observation period belong to Class I, patients with the highest HAQ at baseline – and the group most likely to represent patients eligible for biologics in the UK. It is unclear from the report how much of the ERAS dataset belongs to Class I nor what the rate of drop out is for this class specifically. But AbbVie would contend that the sample size is probably relatively small and the drop-out rate higher and steeper than any of the other latent classes, rendering any NMAR techniques almost void if more than 90% of the sample have missing observations from early on in the observation period.

1.3 Focus on the latent class growth model using the ERAS dataset and sub-analysis of patients after failing two DMARDs (section 4.4 of DSU report)

1.3.1 General issues with the LCGM based on ERAS data used to inform HAQ progression in AG model

The covariates used to predict latent class membership in the DSU report are the same as those employed by Norton et al and are as follows: age, gender, DAS, symptom duration, rheumatoid factor, and fulfilment of ACR criteria, all at baseline. Probably the most frustrating aspect of the latent class model is that it is only the starting criteria that affect a patients' probability of belonging to a class, and the starting characteristics are taken from a cohort of early RA patients (disease <2 years) who were DMARD naïve – i.e. "*explanatory variables influence class membership but not HAQ trajectory*" (page 21 of DSU report). This means that baseline characteristics will affect which class a given RA patient will belong to, but the HAQ trajectories are plotted for each class based on the actual observed data from the cohort of patients belonging to those classes at baseline, irrespective of drop-out. Furthermore, this analysis relies on the assumption that the patient characteristics identified above adequately capture disease progression in the relevant patient population. This is where AbbVie contends the analysis is inherently flawed.

Given that only 26% of patients in the ERAS cohort failed two DMARDs during the entire observation period, it would suggest that the remaining 74% are responding to DMARD therapy or do not require it(!). In which case the HAQ observations over time will increasingly apply to a DMARD responder RA population and it is therefore not surprising that HAQ progression slows. In addition, and as mentioned in 1.2 above, only 10% of patients in the ERAS cohort had high disease activity and failed 2 DMARDs, the criteria required by NICE to receive a biologic in the UK. Problems with the HAQ observations over time are further compounded by the drop-out rate – for example at the end of the observation period, t=15 years only 2.45% of the original sample remain, and indeed from around 10 years since the start of the cohort less than 20% of the original sample remain to contribute observations.

Interestingly, symptom duration at baseline affects latent class membership - yet ERAS is an inception cohort where the average symptom duration is 8.16 months. This is in comparison to some of the RCTs in DMARD failure RA patients for the earlier biologics like adalimumab, etanercept and infliximab where disease duration prior to initiating a biologic is around 11-12 years. Therefore, similarly to the impact DMARD failure has on HAQ progression discussed in 1.2.2., AbbVie would argue that symptom duration measured in months as a predictor for HAQ progression over a lifetime (30-40 years) is not clinically plausible, particularly with the level of drop-out already discussed. The ERAS dataset has less than 20% of the original sample remaining (10-11 years) to inform the HAQ observations at the point at which data for some of the biologic RCTs becomes relevant. At 15 years, approximately 2.5% of the overall sample remains to inform the HAQ progression rates, let alone imagining what the sample sizes are in each of the four latent classes at this time point. Yet Leigh et al showed that patients with disease duration > 20 years have worse HAQ progression rates than those with less than 20 years of disease. If the data reported by Leigh et al are considered to be of merit, then the HAQ trajectories predicted by the ERAS dataset will grossly underestimate the HAQ progression rate in cDMARD failure patients with moderate to severely active RA based on the current analysis.

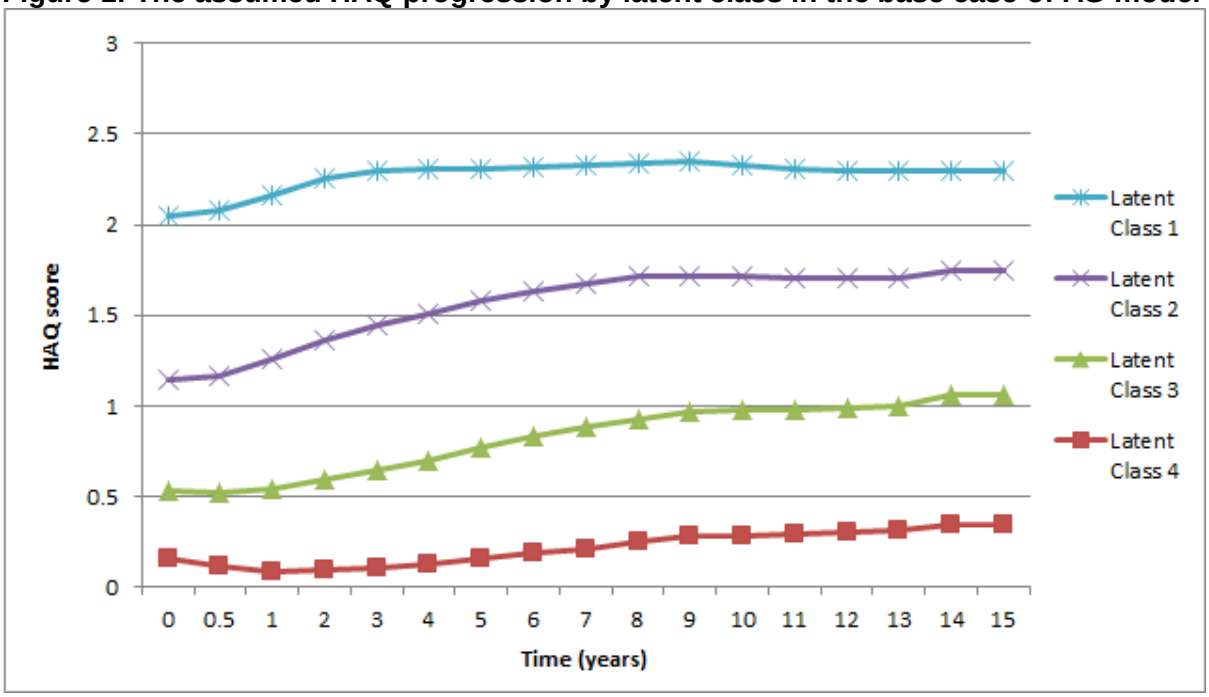
On page 51 of the DSU report the authors acknowledge the issue AbbVie has pointed out above with disease duration: *"Within the ERAS data there is a negative correlation between disease duration and baseline HAQ. This stands to reason as one might expect that those with more aggressive disease, manifesting itself in functional disability, would have a more rapid worsening of their symptoms causing them to seek help at an early time than the average early RA patient, and therefore receiving a first visit with a rheumatologist earlier. However, this is a relationship observed within the small range of disease durations seen amongst the ERAS baseline population. Once this is applied to disease durations exceeding that by a factor of ten this relationship is questionable and is a substantial contributory factor in the predicted HAQ reducing at longer time periods. This demonstrates a limitation of extrapolation beyond the data."*

The DSU report categorically states that this assumption of negative correlation with disease duration and HAQ progression is a substantial contributory factor in the predicted HAQ reducing at longer time periods. Yet, the DSU notes that this relationship is more than questionable, particularly when the evidence regarding HAQ data for patients with disease durations >10 years prior to receiving a biologic is taken into account. AbbVie concludes that the assumptions underpinning the reducing HAQ rate over long periods of time are flawed and therefore all the cost-effectiveness analyses in the AG model that are based on the HAQ progression of cDMARDs using the LCGM based on the ERAS dataset, should be disregarded. The LCGM based on the ERAS dataset, whilst mathematically elegant, does not accurately predict HAQ progression in the relevant RA population.

1.3.2 Implementing a J shaped curve for HAQ progression in cDMARD failure patients - specific impact on moderate RA patients failing cDMARDs

The AG utilised data shown in Figure 1 below in the base case to predict HAQ progression rates for patients on cDMARDs (page 5 of the addendum to the AG report assessing impact of DSU report on cost/QALY gained). As one would expect, based on the baseline criteria applied to assess latent class membership, the majority of patients with severe RA belong to latent classes 1 and 2. However in a post 2009 analysis of the BSRBR baseline characteristics conducted by the AG to reflect current clinical practice, 18% of patients with moderate RA (DAS28 > 3.2) who have failed two DAMRDs were deemed to belong to the bottom two classes (Latent classes 3 and 4).

Figure 1: The assumed HAQ progression by latent class in the base case of AG model



There a couple of additional problems to the issues already highlighted in this document with the LCGM based on the ERAS dataset that have a specific impact on the cost-effectiveness ratios for biologics vs cDMARDs in moderate RA.

Firstly, because ERAS is an inception cohort, patients were cDMARD naïve at baseline. During the first few months of disease, patients would have received either one or two DMARDs either sequentially or in combination depending on treatment practice at the time. As is expected, patients will initially have a good response to cDMARDs, and for some patients they will maintain a good response to cDMARDs – this is particularly obvious in latent classes 3 and 4 – giving rise to the J curve. The populations of interest in this MTA review are populations 2 and 3 defined by the AG – that is severe RA defined by DAS28 >5.1 in patients who have failed 2 cDMARDs and moderate-to-severe RA defined by DAS28 >3.2 in patients who have failed 2 cDMARDs. The definitive criteria being active disease despite treatment with at least 2 cDMARDs. Therefore, using the ERAS inception cohort and modelling an initial response. i.e. an improvement in HAQ to cDMARDs and subsequently applying this improvement to a proportion of moderate or severe RA patients who have already failed 2 DMARDs is clinically and methodologically unsound. Furthermore, it leads to the underestimation of the relative effect of biologics vs. cDMARDs in a DMARD failure population.

Secondly, on page 66 of the DSU report the authors' note, "*We also assume in implementation of these results in the CE model, <u>that HAQ does not continue to rise beyond 15 years for the lowest two latent classes</u>. This makes no difference to the severe*

populations because the probability of being in these low HAQ classes is very small. For the moderate DAS group this assumption may be more relevant because there is a 0.18 probability of being in either class." Even with all the limitations associated with ERAS in terms of the relevant RA population, the high attrition rate and the lack of patients failing cDMARD treatment, assuming that patients in either latent class 3 or 4 have zero progression rate after the 15 year observation period is completely unfounded. This means that the AG are assuming for a proportion of patients receiving cDMARDs that between the time-point of 15 years from entry into the model to the point at which the patient exits the model (i.e. death), which is around 25-30 years, that there is zero progression of their disease. This is an extraordinary assumption to make for the populations of interest in this review who have active progressive disease despite cDMARD treatment.

As a consequence of allowing cDMARD failures to have an initial improvement in their HAQ scores and then preventing any disease progression over a 30 year period (i.e. from t=15 years to death), the effectiveness of cDMARDs for a proportion of patients in the model has been grossly over-estimated. This has obvious implications for calculations of the relative effectiveness of biologics vs. cDMARDs, and on the subsequent ICERS of the biologics vs. cDMARDs in patients with moderate-to-severe RA. AbbVie contends that if the initial improvement applied to the cDMARD arm is removed, progression rates are not capped at 15 years, and a more appropriate rate is used to model HAQ progression, then the ICERs for biologics vs. conventional DMARDs in patients with moderate RA would fall within acceptable cost-effectiveness thresholds as per the manufacturer models submitted as part of this MTA.

1.3.3 Sub-group analysis of patients failing 2 DMARDs

On page 52 of the DSU report, the authors note that all analyses used in the Assessment Group model and the first part of the DSU report include "*observations from patients at different stages of disease and with different characteristics to those that are eligible, or potentially eligible, for biologic drug treatment under NICE guidelines*". The authors justify the inclusion of the entire datasets because "*they attempted to adjust for these differences within the modelling*." However, AbbVie contends that these differences have not been adjusted for, as only a few characteristics at baseline affect membership to specific latent class, they do not inform the HAQ trajectories. Instead the data are taken from the observed data irrespective of DMARD failure status, persistent active disease despite cDMARD treatment, disease duration, bias introduced by the high drop-out rate etc. (issues already discussed). AbbVie considers that the DSU's sub-group analysis (Section 4.4 of the DSU report) attempting to model HAQ trajectories in patients drawn solely from the relevant (or at least, more relevant) groups of patients, whilst not without its limitations, is much more appropriate.

Unfortunately, limitations associated with the cohorts meant that the DSU couldn't do analyses for patients with DAS>5.1 who had failed 2 DMARDs, or even for those with DAS28 >3.2, but it was possible for the DSU to conduct an analysis using HAQ observations for patients who had failed at least 2 DMARDs where t=0 was set at the point a third DMARD was initiated. Sub-group analyses examining HAQ progression rates in these patients were conducted for the ERAS, ERAN and NDB cohorts. AbbVie's response focuses on the ERAS dataset as it was this cohort in its entirety that was used to inform the LCGM in the AG's economic model.

Table 24 on page 55 of the DSU report shows that the mean rate of change from t=2 to the end of the observation period in the ERAS dataset for patients initiating a third DMARD (i.e. 2 DMARD failure patients) was 0.0412 – this is very close to the annual HAQ rate of 0.045 in the manufacturer models and the previous NICE appraisals. Furthermore, the 0.0412 annual

HAQ progression rate from ERAS includes all patients failing 2 DMARDs and not just those with severe disease (DAS28 score > 5.1), which the DSU acknowledges only constitutes 25% of the sub-group failing 2 DMARDs in ERAS. It is therefore highly likely that if it was possible to isolate moderate or severe disease from the sub-group failing 2 DMARDs in ERAS (n=380 at baseline), then the average annual HAQ progression rate would be higher than 0.045 assumed in the manufacturer models and previous NICE appraisals.

AbbVie considers there are additional parameters to the assumed HAQ progression rate for cDMARDs that adversely affect the cost-effectiveness of biologics vs cDMARDs in the Assessment Group's economic model that were pointed out in the 23 September 2013 response document to the first AG report. However for the purpose of this consultation, assuming a linear HAQ progression rate of 0.045 for cDMARDs and 0.06 rate for palliation and utilising the mapping algorithm from Malottki (Table 184, Page 406 in the AG report), the ICERs for the biologics vs. conventional DMARDs are £30,000-33,000/QALY gained. If, as the mean annual HAQ rate for DMARD failure patients with moderate or severe disease from the ERAS data suggests, the annual HAQ progression rate is worse than the rate assumed for the DSA in Table 184, then the ICERs for the biologics vs. conventional DMARDs would fall within acceptable thresholds of cost-effectiveness. This is in line with all the cost-effectiveness estimates for the individual biologics vs. cDMARDs presented in each of the manufacturer submissions.

As well as presenting the average annual rate of change for HAQ observations from ERAS for the DMARD failure sub-group in ERAS, the DSU used latent class growth analysis to analyse the data. As discussed previously, the LCG model allows patient characteristics to influence class membership but does not influence the subsequent HAQ trajectories. The patient characteristics which were allowed to influence class membership for the sub-group analysis were age, DAS and disease duration. Even with the limitations discussed in Sections 1.1 and 1.2 above, the quadratic LCGA using the sub-group of DMARD failures from the ERAS dataset estimates the annual rate of HAQ change to be about 0.0406, with a worse rate estimated from t=2 to t=8 (Table 27, page 58 of the DSU report).

It should be noted that none of the manufacturer submissions model a worse rate than 0.045 for HAQ progression between years 2-8 on cDMARDs, as some of the evidence suggests. Therefore, AbbVie would argue that the estimates of cost-effectiveness presented in AbbVie's original submission, and indeed supported by the other manufacturer submissions, are conservative in this regard but more in line with the evidence for the most likely HAQ progression in patients with moderate and severe active disease who have failed 2 DMARDs, circa £16,500/QALY gained and £19,000/QALY gained, both vs. cDMARDs, for severe (DAS28>5.1) and moderate (DAS28>3.2) RA respectively.

The DSU discusses its conclusions in Section 5 of the report, and concludes that whilst the sub-group analyses in patients who have failed two DMARDs results in a much reduced sample size, it does "*allow more meaningful comparisons to be made*" (page 69 of DSU report). Furthermore, and most importantly the DSU notes for the subgroup analysis of ERA that "*the latent class analysis does differ from the analyses conducted using the full dataset in suggesting the rate of HAQ continues to rise, albeit at a slower rate, particularly for those in the highest latent class. Overall these rates still suggest a lower overall rate of worsening than 0.045 per annum though the predictions for the severe disease subgroup are very close to this."* This conclusion is basically stating that when the relevant RA population is studied i.e. severe active disease despite treatment with 2 DMARDs, the annual rate of HAQ progression is exactly what has been assumed in all the manufacturer models and all previous NICE appraisals.

It can also be inferred that the concept of a LCGM based on the entire ERAS cohort, but then 'adjusting' for different characteristics, to estimate HAQ progression is inherently flawed as AbbVie has argued in Sections 1.1 and 1.3 above. If the HAQ trajectories predicted using the entire ERAS dataset weren't flawed, then one would expect the predictions for HAQ progression using the entire cohort - but making adjustments for disease characteristics - to produce similar results to the HAQ progression rates determined when the relevant population was examined separately.

2.0 AbbVie comments on addendum to the AG report:

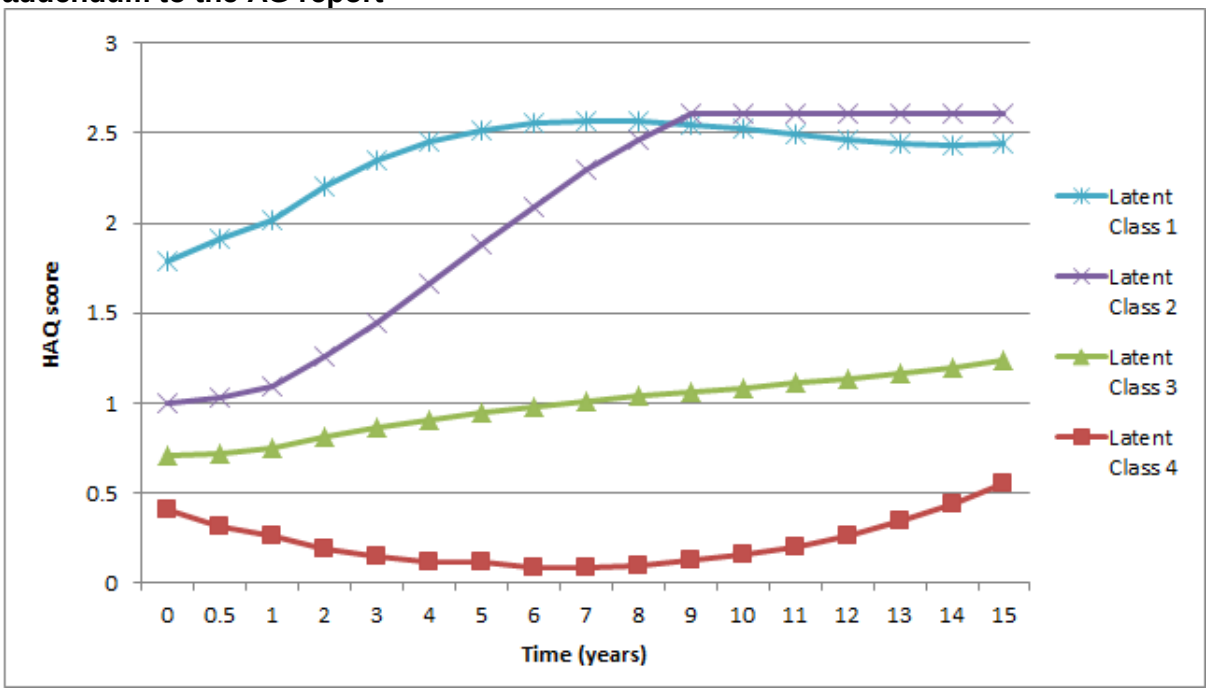
The addendum to the AG report under consultation aims to assess the sensitivity of the cost per QALY gained for biologics to scenarios suggested by the DSU report "*that were deemed worthy of investigation*."

On page 4 of the addendum, the AG note that "*The findings of the DSU report do not support the Assessment Group altering its base case scenario, and as such the results presented in the February 2015 report remain the Assessment Group's best estimation of the likely ICERs of bDMARDs".* AbbVie contends that the findings from the DSU report do in fact suggest that the AG should alter its base case scenario. All the issues raised in Sections 1.1-1.3 above support the use of an alternative HAQ progression rate for cDMARDs than the latent class model adapted from Norton et al and based on the entire ERAS dataset. Whilst AbbVie acknowledges that there is never likely to be the ideal evidence base to estimate the true rate of HAQ progression in cDMARD failures, the sub-group analyses conducted by the DSU suggest that the actual progression rate for the 'relevant' RA population is much more likely to be close to the 0.045 rate assumed in all the manufacturer models and in previous NICE appraisals.

In the addendum to the AG report, the Assessment Group only explore alternative ICERs for patients within each of the four previously defined latent classes from the entire ERAS cohort in whom HAQ increases most in the 15 years from initiation of cDMARDs utilising the Roy-Muthen method to account for data not missing at random (NMAR) discussed in Section 4.5.4 of the DSU report (and discussed in 1.2.4 above). On page 62 of the DSU report, the DSU argues that using the Roy-Muthen method, "*the model identifies, within each latent trajectory class, three sub-classes based on the latent dropouts. These classes show the predicted course of HAQ over time had the patients not dropped out. Table 28 shows that there is a strong relationship between membership of these three classes and the time of dropout. Those that drop out early are likely to be in class 1. For those with the lowest starting HAQ these early dropouts are predicted to move to a very low HAQ. This group seems compatible with those that dropout due to disease remission. Class 1 in the two high HAQ latent classes have the highest rate of worsening in HAQ, which then flattens over time. These fast worsening groups may be those that withdraw due to death or simply because of disease severity."* This is supported by the CORRONA analysis AbbVie conducted and mentioned in 1.2.3, which shows that up to 32.0% of RA patients switched from cDMARDs to a biologic between years 3 to 5 from t=0 due to active progressive disease.

The Assessment Group postulates that patients with the worse HAQ progression over the 15 year period of the ERAS cohort can be captured by the credible upper bounds from the HAQ trajectories estimated using the Roy-Muthen method for each of the latent classes based on the entire ERAS cohort. The AG then calls these patients "*patients with the greatest HAQ progression"* and utilises the data to estimate the plausible lower bound for the ICERs if it was possible to identify such patients. The HAQ progression used in the model for patients with the greatest HAQ progression in each of the four latent classes is shown in Figure 2 below.

Figure 2: The assumed HAQ progression by latent class in those patients with the greatest HAQ progression assumed by the AG in its exploratory analyses in the addendum to the AG report



On page 4 of the addendum, AbbVie notes that for all the exploratory analyses conducted by the AG to investigate a plausible lower bound for the ICERs the AG states *"For all latent classes it was assumed that HAQ would be flat beyond 15 years. This assumption is supported by the data for latent Classes 1 and 2, where the majority of patients reside, there is no indication of a trend for increasing HAQ between 10 and 15 years".* This is an extraordinary assumption for the AG to make, and not supported by the DSU report, which does something similar but only for the latent classes 3 and 4. AbbVie has already argued in section 1.3.2 the implausibility of such an assumption. But in this case, the impact of the AG's assumption is even worse than discussed in section 1.3.2 because it has been applied to all the latent classes including 1 and 2 - the classes containing all the severe RA population estimated by the AG.

AbbVie contends that as a result of this assumption i.e. for all latent classes HAQ is assumed to have zero progression beyond 15 years, the entire exploratory analyses to investigate the plausible lower bound ICERs in the addendum to the AG report are rendered completely void, inappropriate, and as a result underestimate the cost-effectiveness of the biologics vs cDMARDs.

Firstly and as discussed previously, the credible upper bounds driving the trajectories in Figure 2 above are based on NMAR assumptions for a dataset that has severe limitations in terms of actually relating to the relevant RA population (discussed in more detail in Sections 11.1 and 1.2 above) – i.e.

- it includes mild and self-limiting RA disease, where some patients will go in to remission - a population not relevant to this appraisal;
- 75% of patients have not failed cDMARDs and are essentially still cDMARD responders at the end of the 15 year observation period;
- the attrition rate is so high, the only patients remaining in the cohort beyond 10 years (less than 20% remaining) are a self-selected group of cDMARD responders, which is why the HAQ progression curve flattens.

Secondly, results from the Roy-Muthen method showed that those patients that are likely to drop out early in the observation period belong to Class I, patients with the highest HAQ at baseline – and the group most likely to represent patients eligible for biologics in the UK. It is unclear from the report how much of the ERAS dataset belongs to Class I nor what the rate of drop out is for this class specifically. But AbbVie would contend that the sample size is probably relatively small and the drop-out rate higher and steeper than any of the other latent classes, rendering any NMAR techniques almost void if more than 90% of the sample have missing observations from early on in the observation period. Such that the NMAR estimates informing the credible limit from t=10-15 years, which in the DSU's analysis suggest HAQ progression is zero during this 5 year observation period, will be based on a few cDMARD responders left in that class.

Finally, from a clinical point of view, assuming that <u>all</u> patients from the ERAS dataset have zero HAQ progression after the 15 year observation period is completely implausible. Whilst the entire dataset is not representative of RA patients eligible for a biologic in the UK, it will contain a proportion of patients who will fail 2 or more cDMARDs and have active severe disease during the observation period i.e. some patients will have characteristics similar to the established RA population observed in the RCTs for the earlier biologics: adalimumab, etanercepy, and infliximab e.g. average of 11 years disease duration, average DAS28 at baseline of 6.8 and average no. DMARDs failed = 3. To suggest that these patients with active progressive, debilitating disease have zero HAQ progression is completely counterintuitive to the evidence base.

As discussed in 1.3.2, the impact of assuming zero progression beyond 15 years for patients receiving cDMARDs has a considerable impact on the ICERs for biologics vs. cDMARDs. To put it simplistically, what the AG has assumed is that for a given patient irrespective of their potential for failing DMARDs and any other prognostic factors beyond some baseline characteristics, that they enter the model at t=0 and follow one of the four latent classes to time-point t=15 years, from this point up until the point at which the patient exits the model (i.e. death), which is around 25-30 years, the AG has assumed that there is <u>zero</u> progression of their disease. So for 25-30 years of the time horizon within the model all cDMARD failure patients are assumed to have zero progression for a known, progressive and debilitating, destructive disease. This is an extraordinary assumption to make.

As a result of this, any other discussion around the appropriateness of the AG's exploratory analyses isn't necessary because the main parameters underpinning the analyses are flawed. So where the AG argues that ICERs circa £40,000 per QALY gained for the biologics vs cDMARDs are the most optimistic estimates i.e. the most plausible lower bound for the ICERs – is wholly incorrect. Because the AG artificially flattens the HAQ progression rate beyond 15 years for cDMARDs, the relative effectiveness of biologics vs. cDMARDs are underestimated with the resultant impact being that the ICERs for the biologics vs. cDMARDs are artificially high. AbbVie contends that if a more accurate rate of HAQ progression was used and progression for all classes wasn't capped at 15 years, then the ICERs for the biologics vs. conventional DMARDs would be well within accepted cost-effectiveness thresholds – and in line with AbbVie's own estimates for the cost-effectiveness of biologics vs. cDMARDs in moderate and severe RA, i.e. under £20,000/QALY gained.

3.0 Overall conclusion:

It is clear from the comprehensive analysis undertaken by the DSU that the true estimates for HAQ progression in cDMARD failure patients can never be determined. This is not surprising, it would be extremely difficult to justify leaving an RA patient with active disease despite treatment with 2 or more cDMARDs in chronic, debilitating pain incurring progressive

disability for any period of time, let alone for a period of 15 years, when there are extremely efficacious treatment available (i.e. biologics).

AbbVie considers that the latent class growth model to predict HAQ trajectories based on the entire ERAS dataset is not a suitable source to model HAQ progression. In the absence of better estimates, AbbVie suggests that the output from the DSU's sub-group analysis in patients who have failed 2 DMARDs is more appropriate to model the most likely HAQ progression rates in population failing cDMARDs. There is a considerable amount of data showing that HAQ progression in patients responding to a biologic is zero. It would therefore be unfair to penalise patients requiring biologics because the true HAQ progression rate for a cDMARD failure population can't be estimated effectively. Whilst AbbVie understands that the rate of HAQ progression of cDMARDs is a large driver in the estimates of cost-effectiveness, the uncertainty around the true rate is never going to be resolved as it will never be possible to capture HAQ progression in patients failing cDMARDs over a long period of time in clinical practice. In which case, in the absence of robust data in the relevant RA patient population, the Appraisal Committee needs to consider other factors that have not been captured in the modelling that mitigate some of this uncertainty. For example the considerable impact biologics have on societal costs, which are not included, as part of a wider value judgement would considerably lower the ICERs of the biologics vs. cDMARDs.