Alemtuzumab for the treatment of relapsing-remitting multiple sclerosis

ERG overview of manufacturer's response to the ACD

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors Emma Loveman

Jeremy Jones

Jackie Bryant

Keith Cooper

Correspondence to Emma Loveman

Senior Research Fellow

Southampton Health Technology Assessments Centre (SHTAC)

University of Southampton

First Floor, Epsilon House

Enterprise Road, Southampton Science Park

Southampton SO16 7NS

Date completed 04/02/2014

Introduction

The manufacturer of alemtuzumab provided a document outlining their response to the ACD.

SHTAC was requested to provide an independent assessment of the two subgroup analyses submitted by the manufacturer in this document. This document refers to the original manufacturer's submission (GMS), the subgroup analyses in response to the ACD (subgroup submission), the original ERG report (ERG report).

ERG critique of Genzyme's subgroup submission (presented as sections A2 and A3).

A2- HARRMS despite treatment subgroup

Clinical trial evidence

The manufacturer submitted evidence for a subgroup with HARRMS despite interferon treatment from the CARE-MS II trial. This was a subgroup noted in the original NICE scope. It is presented to allow an assessment of the use of alemtuzumab compared with fingolimod (a NICE scoped comparator). The ERG clinical advisors suggest that alemtuzumab may be used as an alternative to fingolimod for highly active RRMS, however, is unlikely to replace it entirely. One ERG advisor also comments that neurologists may also wish to use alemtuzumab as a third line after treatment with fingolimod has failed, but this is not relevant to the scope.

The subgroup was defined as: >365 days IFNβ-1a in the year prior to treatment and gadolinium enhancing lesion positive or T2 > 0.5ml and 1 or more relapses in prior year. In the original GMS, and in the Clinical Study Report, the criteria for highly active RRMS was reported to be 'those who experienced 2 or more relapses in the year prior to randomisation and had at least 1 gadolinium-enhancing lesion at baseline' (without the addition of 'or T2 > 0.5ml). The proportion of trial participants meeting these criteria was reported to be 53%.

The data presented by the manufacturer were not previously reported in the GMS. The GMS and CSRs present subgroup analyses for time to SAD (years 0 to 2) and relapse rate for those described as 'highly active' (and meeting the criteria noted previously in the GMS) but the numbers of participants in these groups does not correspond to the subgroup sample size presented in the subgroup submission document. The ERG therefore believes that the subgroup data presented have been analysed post-hoc, taking into account the additional criteria of T2 lesion >0.5ml. Caution is therefore recommended in the interpretation of the results.

The ERG is unable to fully comment on the accuracy of the data presented for the subgroup analyses for 6-month SAD, 3-month SAD and ARR. As seen in Table A.2.1.1 the trial subgroup analysis is seen to be favourable to alemtuzumab versus IFNβ-1a for 6 month SAD. This is in line with the direction of effect seen in the full data set (presented in the subgroup submission and the GMS). Table A2.1.2 shows that the trial subgroup analysis crosses the line of no effect for the 3 month SAD.

[redacted] Table A2.1.3 shows that the trial subgroup analysis is favourable to alemtuzumab for ARR. This appears to be in line with the full data set presented, although slightly different from the GMS values.

By observation the manufacturer states that data available from the fingolimod RCT (3 month SAD and ARR) are comparable to the alemtuzumab data. The relapse definitions were slightly different between the two trials. The ERG clinical advisors state that while there are subtle differences between these groups, they are not clinically significant and would be the same people in practice. These differences reflect the slight variability in prescribing of current drugs by MS Neurologists. One advisor comments, however, that non-response to treatment and highly active disease may reflect two different things. That relapse in 12 months may be high, but it could be years later before another relapse occurs.

MTC evidence

The MTC evidence presented in tables A.2.2.1.1 and A.2.2.2.1 is identical to that previously presented in the GMS (see Table 6.7.13 and 6.7.14, page 171 of the GMS). As previously noted in the ERG report (page 35), the evidence networks presented in Figures A2.2.1.1 and A2.2.2.1 appear inconsistent in their approach to multi-arm trials within networks (including three arms from FREEDOMS (two doses of fingolimod and placebo), but only two arms from CARE MS II (excluding 24mg alemtuzumab) for 3 month SAD) and between analyses (including two arms from FREEDOMS (one dose of fingolimod and placebo) for ARR vs three arms for 3 month SAD). It is unlikely that this will have a substantial impact on the results of the MTC but it is not clear why this approach has been adopted or whether an adjustment has been made in the analysis for the inclusion of multi-arm studies (see ERG report, pages 31 and 34 for comments on original analyses of the full data set). The evidence network is highly dependent on the teriflunomide trials (TEMSO, TOWER and TENERE) – it should be noted that the data from TENERE trial are for all patients and not the sub-group being considered (see Page 169 of the GMS), while data from TOWER and TEMSO are for a sub-group of previously-treated patients which the manufacturer refers to

as a “proxy for HA RRMS despite interferon use”. This constitutes a substantial weakness in the analysis and needs to be considered as an additional source of uncertainty in interpreting the results of the MTC. The limitations of the data are stated by the manufacturer in the assessment of clinical heterogeneity.

The absence of both indirect and direct evidence for links in the networks means that formal assessment of inconsistency in the network could not be carried out. However the manufacturer could have considered tabulating results for all comparisons in the TEMSO, TOWER and TENERE trials (beyond those reported in Table A2.2.2.2) to support an informal assessment of consistency in these trials which are key to considering the validity of the MTC results.

The analysis of SAD is limited to the measure at 3 months, as no data could be found for 6 month SAD for fingolimod. This limitation is recognised by the manufacturer (see bottom of page 8 of the subgroup submission document).

A3 - RES RRMS in treatment naive or previously treated subgroup

Clinical trial evidence

The manufacturer submitted evidence for a subgroup with RES RRMS in treatment naive or those previously treated (data from CAMMS223, CARE-MS I and CARE-MS II trials). This was a subgroup noted in the original NICE scope. It is presented to allow an assessment of the use of alemtuzumab compared with natalizumab (a NICE scoped comparator). The ERG clinical advisors suggest that alemtuzumab would be considered as an alternative to nataluzimab for RES RRMS. For some patients it is unlikely to replace natalizumab however. One ERG advisor states that alemtuzumab may be used as a third line after treatment with nataluzimab has failed.

The subgroup was defined as at least two relapses in the year prior to randomization and a gadolinium enhancing lesion positive at baseline. This is the definition reported in the original GMS for the CAMMS223 trial and the CARE-MS-II trial, and in the CSRs for the CARE-MS-I and CARE-MS-II trials (for highly active disease) respectively. No subgroup analyse were presented for CARE MS I in the original GMS. Although subgroup analyses for a highly active group (defined as above; 2 relapses and a GE lesion at baseline) is reported in the CSR there has been no publication of these data and the original GMS appears to only focus on subgroups that were published. It is unclear to the ERG if the subgroup analyses from the CAMMS223 trial is post hoc analysis as it does not appear to be reported in the CSRs. In

the GMS subgroup analyses were reported from CAMMS223 based on a publication by Wingerchuk 2010, however, limited details were reported. The definition for relapse in CAMMS223 is reported in Tables A.3.1.1-3 as ≥ 2 episodes in prior year and GAD+ but in Table A3.2.4.1 it is reported that data on this measure were not collected in the trial. Caution is therefore recommended when considering the results.

The ERG is unable to comment on the accuracy of the data presented for the subgroup analyses for 6-month SAD, 3-month SAD and ARR. The presented data for the full data sets concur with those in the original GMS and the direction of effects seen were mostly similar (in the CARE-MS-II the subgroup analysis of 6-month SAD was no longer statistically significant).

Trial data were meta-analysed but the approach used was not described. Tables A3.1.1 - A3.1.3 show the results of the individual trials and the meta-analysed endpoints for 6-month SAD, 3-month SAD and ARR. The pooled hazard ratios suggest that the 6-month SAD and ARR were favourable to alemtuzumab [redacted] CAMMS223 contributes the most data to this analysis, and as noted above it is unclear whether these data were post-hoc. If CAMMS223 trial data are removed from the meta-analysis the mean effect is likely to move nearer to the line of no effect.

The proportions of participants meeting the criteria for RES RRMS differed between the three trials, in CARE-MS I and II these were 29% and 23% respectively, in the CAMMS223 trial this was 56%. In the pooled analysis the rate was therefore 30%.

By observation the manufacturer states that data available from the natalizumab RCT (3 month SAD and 6 month SAD) are comparable to the alemtuzumab data. The relapse definitions were slightly different between the two trials. Similarly to above, the ERG clinical advisors suggest that while these groups are slightly different in clinical practice they would be the same people. One advisor has commented that limiting RES to those with GAD lesions at baseline may not be reliable as it depends at what time point you perform the scan.

MTC evidence

The MTC evidence presented in tables A.3.2.2.1 and A.3.2.3.1 is identical to that previously presented in the GMS (see Table 6.7.16 and 6.7.17, page 174 of the GMS). The data for six month SAD for this sub-group (results presented in Table A.3.2.1.1) was not presented in

the GMS. As noted above the evidence networks presented in Figures A3.2.2.2 and A3.2.3.1 appear inconsistent in their approach to multi-arm trials within networks and between analyses. The evidence networks also include irrelevant and unconnected (other than to placebo) comparators. This is not expected to have a substantial impact on the MTC results but it is not clear why this approach has been adopted or whether an adjustment has been made in the analysis for the inclusion of multi-arm studies (see ERG report, pages 31 and 34 for comments on original analyses of the full data set).

The evidence networks for the MTC are highly dependent on the TEMSO and TENERE trials to link SC-IFN-1a and placebo, but there is limited discussion of the appropriateness of including these studies for this sub-group of patients. The manufacturer notes that only the AFFIRM study (providing data on natalizumab vs placebo) used a definition of RES RRMS adopted within NICE guidance. A rationale is provided for use of data from the TEMSO trial (in preference to TOWER) based on reporting of GAD lesions on brain MRI, but no rationale for inclusion of TENERE is provided other than it "needed to be included … to allow … alemtuzumab versus natalizumab comparison to be carried out" (p27 subgroup submission document). The data from TENERE is for a sub-group from the trial population who had two or more relapses in the prior year at baseline (but include no information on GAD lesions).

The manufacturer's discussion of clinical heterogeneity focuses primarily on the baseline characteristics of patients in the natalizumab and alemtuzumab trials – it is not clear why data on GAD is missing for CAMMS223 while the definition of RES for drawing sub-groups of patients from this trial requires baseline GAD (in all cases the CAMMS223 trial provide the most favourable effect estimates for alemtuzumab). Inconsistency in definitions of RES used to draw sub-group data from trials providing data for the MTC constitute a serious weakness in the analysis and need to be considered as an additional source of uncertainty in interpreting the results of the MTC. The limitations of the data are stated by the manufacturer in the assessment of clinical heterogeneity, but the weaknesses of the evidence base and additional (methodological) uncertainty around parameter estimates is not acknowledged. It would also appear that gaps in the evidence and inconsistencies between stated criteria and those used in the analysis (i.e definition used to drawing RES sub-group from CAMMS223) are not fully acknowledged in the subgroup submission.

The absence of both indirect and direct evidence for links in the networks means that formal assessment of inconsistency in the network could not be carried out.

The manufacturer states in the first bullet point under A.3.2.3 (results of ARR MTC) that alemtuzumab is [redacted] (presumably based on the credible range (labelled CI) excluding 1), but fails to note that the same applies for natalizumab. In addition, it is not clear how ARR data might be included in the MTC for trials where Table A.3.2.3.2 states “ARR RR not reported” – for example, the ERG is unclear how a result from the MTC for fingolimod 0.5mg vs placebo can be included in table A.3.2.3.1 when then entry for fingolimod 0.5mg vs placebo (the only link for fingolimod 0.5mg into the evidence network) in Table A.3.2.3.2 (ARR RRs of data input into ARR RR MTC) states “ARR RR not reported”. Fingolimod (which may be regarded as an irrelevant comparator for this sub-group) appears to have been included in the evidence network as a link between IM-IFNβ-1a and placebo.

Overall the manufacturer’s narrative for this sub-group has a tendency to over-interpret statistical results. For example, the statement “Alemtuzumab is [redacted]” is made where there is a wide credible interval. This narrative underplays large potential sources of clinical heterogeneity as well as substantial uncertainty arising from the weaknesses of the evidence base. There is additional methodological uncertainty arising from studies included in the MTC (more truly a chain of indirect comparisons) to link comparators of interest. The ERG have further concerns regarding the definitions of RES and whether the data drawn from included trials have been consistent with the stated definitions.

Summary

Overall the ERG note that while these data provide some evidence of the effectiveness of alemtuzumab in the two reported subgroups, there are a number of uncertainties with these data and results need to be cautiously interpreted.