# Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

# Alemtuzumab for the treatment of relapsing-remitting multiple sclerosis


**Produced by** Southampton Health Technology Assessments Centre

**Authors** Keith Cooper
Jackie Bryant
Petra Harris
Emma Loveman
Jeremy Jones
Karen Welch

**Correspondence to** Emma Loveman
Southampton Health Technology Assessments Centre
University of Southampton
First Floor, Epsilon House
Enterprise Road, Southampton Science Park
Southampton SO16 7NS



**Date completed** 24th September 2013

**Source of funding**: This report was commissioned by the NIHR HTA Programme as project number 12/83/01.

**Declared competing interests of the authors**

None

**Acknowledgements**

We are very grateful to the following experts who offered clinical advice and comments on the draft report: Dr Sarah Cader, Consultant Neurologist (North Hampshire Hospital and Wessex Neurological centre, Southampton) and Dr Ben Turner, Consultant Neurologist and Honorary Senior Lecturer (Barts and The London NHS Trust).

We also thank Dr Jill Colquitt, Senior Research Fellow, SHTAC, for acting as internal editor for the ERG report and Mrs Neelam Kalita, Research Fellow, SHTAC, for reviewing an earlier version of this report.

**Rider on responsibility for report**

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

**This report should be referenced as follows:**

Cooper, K, Bryant J, Harris P, Loveman E, Jones J, Welch K. Alemtuzumab for the treatment of relapsing-remitting multiple sclerosis: A Single Technology Appraisal. SHTAC, 2013.

**Contributions of authors**

K Cooper (Senior Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; J Bryant (Principal Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; P Harris (Research Fellow) critically appraised the clinical effectiveness systematic review, the mixed treatment comparison and drafted the report; E Loveman (Senior Research Fellow) critically appraised the clinical effectiveness systematic review, drafted the report and project managed the review; J Jones (Principal Research Fellow) critically appraised the mixed treatment comparison and drafted the report; K Welch (Information Specialist) critically appraised the search strategy and drafted the report.

**Word count:** 30,305

[redacted].

TABLE OF CONTENTS

1 Introduction to ERG Report ....11
2 BACKGROUND ....11
2.1 Critique of manufacturer's description of underlying health problem ....11
2.2 Critique of manufacturer's overview of current service provision ....11
2.3 Critique of manufacturer's definition of decision problem ....12
3 CLINICAL EFFECTIVENESS ....13
3.1 Critique of manufacturer's approach to systematic review ....13
3.2 Summary statement of manufacturer's approach to evidence synthesis ....35
3.3 Summary of submitted evidence ....36
4 ECONOMIC EVALUATION ....51
4.1 Overview of manufacturer's economic evaluation ....51
4.2 Critical appraisal of the manufacturer's submitted economic evaluation ....54
4.3 Additional work undertaken by the ERG ....75
4.4 Summary of uncertainties and issues ....81
5 End of life ....83
6 Innovation ....83
7 DISCUSSION ....83
7.1 Summary of clinical effectiveness issues ....83
7.2 Summary of cost effectiveness issues ....83
8 REFERENCES ....84

LIST OF TABLES

Table 1: Main characteristics of participants in the three included RCTs ....21
Table 2 Potential additional studies identified by the ERG ....22
Table 3: Manufacturer and ERG assessment of trial quality ....24
Table 4: ERG appraisal of MTC approach ....30
Table 5 Comparison of three month SAD hazard ratios from base case and all trial MTC versus direct evidence ....33
Table 6 Three month SAD HRs from MTC – ERG analysis ....34
Table 7 Quality assessment (CRD criteria) of MS review ....35
Table 8: Sustained accumulation of disability at 3 months ....36
Table 9: Sustained accumulation of disability at 3 months from the MTC ....37
Table 10: Sustained accumulation of disability at 6 months ....37
Table 11: Sustained accumulation of disability at 6 months from the MTC ....38
Table 12: Relapse rate, annualised relapse rate, and proportion relapse free ....39
Table 13: ARR from the MTC ....40
Table 14: Proportion relapse-free from the MTC ....41
Table 15: change in mean EDSS score ....41
Table 16: Overview of AEs in alemtuzumab-treated patients ....45
Table 17: Grouped AEs reported at a ≥5% incidence in alemtuzumab-treated patients ....46
Table 18: Discontinuations due to adverse events odds ratio results from the MTC ....46

Table 19: Thyroid Disorders in Alemtuzumab 12mg -Treated Patients (All Available Follow Up, Pool C) ....48
Table 20: Incidence of Infusion-Associated Reactions by Cycle and Severity in All Alemtuzumab 12mg -Treated Patients (All Available Follow Up, Pool C) ....48
Table 21: First Immune Thrombocytopenic Purpura Event by Cycle in All alemtuzumab 12mg -Treated Patients, All Available Follow Up (Pool C) ....49
Table 22: SAEs rates of IARs, infections, thyroid and ITP per trial ....49
Table 23: Overview of AEs in SC IFNβ-1a-treated patients ....50
Table 24: Overview of SAEs in SC IFNβ-1a-treated patients ....51
Table 25 Base case deterministic cost effectiveness results (GMS B7.7.12 page 338) ....53
Table 26: Critical appraisal checklist of economic evaluation ....55
Table 27 NICE reference case requirements ....55
Table 28: Key clinical effectiveness parameters used in the model ....60
Table 29: EDSS utilities used in the model (from GMS Table B7.4.2) ....65
Table 30: Annual costs associated with the technologies in the economic model (from GMS Table B7.5.7) ....68
Table 31: List of health states and associated costs in the economic model (from GMS Table B7.5.8) ....68
Table 32: Costs associated with relapses included in the economic model ....69
Table 33: One way sensitivity analysis results of alemtuzumab versus SC IFNβ-1a 44 µg in active RRMS (GMS Table B7.7.15) ....71
Table 34: Deterministic parameter sensitivity analysis of alemtuzumab versus SC IFNβ-1a 44µg in active RRMS (GMS Table B7.7.21) ....72
Table 35: Incremental deterministic base-case results of alemtuzumab versus SC IFNβ-1a 44µg ....76
Table 36: ERG Incremental deterministic base-case results of alemtuzumab versus SC IFNβ-1a 44 µg, BSC and glatiramer acetate (with trial characteristics and reduced progression RRMS to SPMS) ....77
Table 37: Incremental deterministic base-case results of alemtuzumab versus SC IFNβ-1a 44 µg for the ERG's preferred base case ....77
Table 38: ERG Subgroup analyses using for treatment naïve and experienced subgroups ....80
Table 39: ERG Subgroup analyses for HA despite interferon use and RES ....81

LIST OF FIGURES

Figure 1: Schematic of the multi-state Markov model ....57
Figure 2: Multi-CEAC of alemtuzumab versus all beta-interferons, glatiramer acetate, fingolimod and natalizumab in active RRMS (GMS Figure B7.7.7) ....74

LIST OF ABBREVIATIONS

| | |
|---|---|
| AE | Adverse event |
| ARR | Annualised relapse rate |
| BSC | Best supportive care |
| CEAC | Cost-effectiveness acceptability curve |
| CHMP | Committee for Medicinal Products for Human Use |
| CI | Confidence interval |
| CrI | Credible Interval |
| CSR | Clinical Study Report |
| DMT | Disease modifying therapy |
| DoH | Department of Health |
| DSA | Deterministic sensitivity analysis |
| DSMB | Data Safety Monitoring Board |
| EDSS | Expanded disability status scale |
| EQ-5D | European Quality of Life-5 Dimensions |
| ERG | Evidence review group |
| GA | Glatiramer acetate |
| GMS | Genzyme manufacturer's submission |
| HR | Hazard ratio |
| HRQL | Health related quality of life |
| ICER | Incremental cost-effectiveness ratio |
| ITT | Intention to treat |
| MS | Multiple sclerosis |
| MTC | Mixed treatment comparison |
| PA | Per annum |
| PAS | Patient access scheme |
| PRMS | Progressive-relapsing multiple sclerosis |
| PSA | Probabilistic sensitivity analysis |
| PSS | Personal social services |
| PSSRU | Personal Social Services Research Unit |
| QALY | Quality adjusted life year |
| RCT | Randomised controlled trial |
| RES | Rapidly evolving severe |
| RR | Relative risk |
| RRMS | Relapsing remitting multiple sclerosis |
| SAD | Sustained accumulation of disability |
| SG | Standard gamble |
| SPMS | Secondary progressive multiple sclerosis |
| TTO | Time-Trade-Off |

## SUMMARY

### Scope of the manufacturer submission

The scope of the submission was to assess the effectiveness and cost-effectiveness of alemtuzumab for the treatment of relapsing remitting multiple sclerosis (RRMS) compared to beta-interferon, glatiramer acetate (GA), natalizumab (for treatment naive or previously treated patients with rapidly evolving severe RRMS) or fingolimod (for patients with highly active RRMS who have been previously treated with beta interferon).

### Summary of submitted clinical effectiveness evidence

The manufacturer's submission to NICE included:

i) a systematic literature review to identify any relevant randomised controlled trials (RCTs) based on any disease modifying treatment (DMT) for relapsing-remitting multiple sclerosis (RRMS) in adults.
ii) two non-RCTs, chosen without a systematic search and not incorporated into a systematic review.
iii) a meta-analysis of the RCTs included in the direct comparison.
iv) a mixed treatment comparison (MTC) based on the literature review for the head-to-head trials. The direct comparisons did not include any trials comparing alemtuzumab with the coped comparators.

The literature review identified 2004 potential trials, of which 52 trials were eligible for qualitative synthesis. Of these, three RCTs included alemtuzumab as a treatment, two trials in treatment-naïve and one trial in treatment-experienced patients. In addition, evidence from two extension studies from the included trials was presented.

Two non-RCTs were included in the report. It is unclear how these were chosen for inclusion in the report and whether any other non-RCTs would be relevant to the decision problem. One of the studies is still unpublished and limited data for both trials were presented in the main report. Due to these reasons the ERG did not assess the non-RCTs.

The meta-analysis was based on three RCTs and combined a patient group consisting of treatment-naïve and treatment-experienced. No subgroup analyses were performed. Patients

Superseded – see Erratum

varied in onset of multiple sclerosis (MS), MS episode history and time since relapse. The ERG does not consider it appropriate to pool these data.

For indirect evidence, the manufacturer included a MTC consisting of 30 RCTs (including the 3 RCTs included in the head-to-head comparison). While no rationale for conducting MTCs was provided, it was acknowledged in discussion that a lack of head-to-head comparisons for included treatment regimens required some form of indirect comparison. Studies in the network were restricted to studies with patient recruitment after 2000 and with ≥80% RRMS patients. Justifications for these restrictions were provided and experts agree that the restrictions are reasonable. Sensitivity analyses were run on datasets without these restrictions.

Treatment with alemtuzumab statistically significantly reduced the rate of relapse when compared with IFNβ-1a in all three included RCTs. The pooled risk ratio (RR) of annualised relapse rate (ARR) for alemtuzumab versus IFNβ-1a from the three RCTs was [redacted] In addition, MTC of alemtuzumab with [redacted] were all statistically significantly different, but not the comparison between alemtuzumab and [redacted].

Sustained accumulation of disability (SAD) at six months was statistically significantly better in those treated with alemtuzumab than those treated with IFNβ-1a in two trials (one in treatment experienced participants, and one in treatment naïve participants). The pooled RR for SAD at 6 months for alemtuzumab versus IFNβ-1a from the three RCTs was [redacted]). In addition, the MTC between alemtuzumab and [redacted] was statistically significantly different.

Some subgroup analyses on a population with rapidly evolving RRMS were undertaken in two trials. Few results were reported. For SAD at six months the hazard ratio favoured treatment with alemtuzumab in one trial in a treatment naive population. No subgroup analyses were undertaken on those with highly active RRMS who had been previously treated with DMTs.

Adverse events were reported by most patients, the incidence of grade 3, 4 and 5 adverse events in all alemtuzumab treated participants (pooled from all three trials) were [redacted] respectively. Adverse event rates were higher in the first year and appeared to decrease over time in most cases, the main exception being thyroid related disorders. Serious

Superseded – see Erratum

adverse event rates declined over time but the rate of SAEs by number of treatment courses received was [redacted] in the alemtuzumab-treated group. [redacted].

**Summary of submitted cost effectiveness evidence**
The manufacturer's submission to NICE includes:

i) a review of published economic evaluations of MS treatments compared for adults with RRMS or progressive MS (including Secondary Progressive MS [SPMS] or Progressive-Relapsing MS [PRMS])

**ii)** a report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of alemtuzumab is compared with beta-interferons, glatiramer acetate, fingolimod and natalizumab for active RRMS.

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of adults with RRMS or progressive MS (including SPMS or PRMS). The review identified 33 studies evaluating cost effectiveness in MS but none of these studies were for alemtuzumab.

The cost effectiveness analysis (CEA) uses a multi-state Markov model to estimate the cost-effectiveness of alemtuzumab compared with other possible alternatives in adults with active RRMS. The model adopted a lifetime horizon of 50 years with a one year cycle length. Modelled health states are based on disease classification (RRMS or SPMS) and severity (defined by the EDSS). The model was based on a structure developed by the School of Health and Related Research (ScHARR) in the evaluation of beta-interferons for the treatment of MS. Active RRMS patients entered the model with baseline characteristics collected for RRMS patients in the UK Risk Sharing Scheme (RSS). Clinical data used in the model are based on results from the alemtuzumab trials and MTC for RRMS (in the base case), HA despite interferon use (in the subgroup analysis) and RES (in the subgroup analysis).

Results from the economic model are presented as incremental cost per QALY gained for alemtuzumab compared with beta interferons, glatiramer acetate, fingolimod and natalizumab. For the base case an incremental cost per QALY gained of £8,924 versus glatiramer acetate is reported. Other analyses show that the comparators were strongly or extendedly dominated.

The model explores structural and parameter uncertainty in one-way and probabilistic sensitivity analyses (PSA). The deterministic sensitivity analysis indicated that the model was most sensitive to the 3-month SAD hazard ratio (HR). The model was also fairly sensitive to the inclusion criteria applied in the derivation of MTC results in terms of SAD, ARR and withdrawal. Scenario analysis showed that the model was also sensitive to assumptions around the waning of treatment effect.

**Commentary on the robustness of submitted evidence**

**Strengths**

The GMS presents reasonably well conducted systematic reviews of evidence of the clinical and cost-effectiveness of alemtuzumab, undertaking a MTC where direct evidence was lacking.

Three RCTs directly comparing alemtuzumab with IFN-β1a used in the submission were reasonably well conducted and provide evidence for the treatment effect in terms of relapse, SAD and adverse events.

The approach taken in the submission to model MS is reasonable and based on previous MS models.

**Weaknesses and Areas of uncertainty**

There were different populations in the three pivotal trials, and the ERG does not consider it was appropriate to pool these data. In one RCT the comparison was with a treatment that many participants had already failed to respond to prior to randomisation.

Data were available to assess treatment naive and previously treated populations in subgroups as per the NICE scope; however, this was not undertaken.

The only head to head comparison was with IFN-β1a. No head to head trials of alemtuzumab with the other comparators have been undertaken and so assessment of effectiveness is based on indirect comparison.

The MTC was conducted reasonably; however, there was limited discussion of the limitations of the analysis, especially with respect to the subgroups analysed.

There are some areas of uncertainty relating to the economic model: there are limitations associated with both methods of estimating disease progression; the choice of studies for informing HRQoL estimates appears arbitrary; and there is uncertainty around the correct value for health state costs.

**Summary of additional work undertaken by the ERG**

The ERG conducted the following additional analyses:

- A preferred base case, with an alternative patient population and different progression rate from RRMS to SPMS;
- A series of sensitivity analyses for this new base case, including varying disease progression by reducing the transition probabilities to more severe health states by 50%; changing quality of life values using upper and lower confidence intervals; varying health state costs; alternative relapse cost of hospitalisation; treatment waning effect; changes in the proportion who receive subsequent doses of alemtuzumab; results from MTC all years ≥80% RRMS; disease progression using 6 month SAD.
- The preferred base case with effectiveness data from the CARE-MS trials for treatment naïve and treatment experienced patients.
- The preferred base case with MTC effectiveness data for the HA despite interferon use and RES subgroups.

In the new base case alemtuzumab is shown to dominate all comparative treatments and be cost effective compared to best supportive care with an ICER of £9907 per QALY gained.

In the sensitivity analyses the ERG found the results robust to changes in assumptions and input parameters, with a cost effectiveness of less than £10,000 per QALY compared to SC IFNβ-1a 44 µg for all analyses.

Subgroup analyses for the HA despite interferon use and RES subgroups show that alemtuzumab continues to dominate fingolimod and natalizumab, respectively.

# 1 Introduction to ERG Report

This report is a critique of the manufacturer's submission to NICE from Genzyme Therapeutics Ltd (referred hereafter as the GMS) on the clinical effectiveness and cost effectiveness of alemtuzumab for the treatment of relapsing-remitting multiple sclerosis (RRMS) in adults. It identifies the strengths and weakness of the GMS. Clinical experts were consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the GMS was requested from the manufacturer by the ERG via NICE on 7th August 2013. A response from the manufacturer via NICE was received by the ERG on 2nd September 2013 and this can be seen in the NICE evaluation report for this appraisal.

# 2 BACKGROUND

## 2.1 Critique of manufacturer's description of underlying health problem

The description of MS appears to be appropriate and outlines the different forms of the disease and the disease course and the ERG's clinical advisors did not have any major concerns with the descriptions provided.

## 2.2 Critique of manufacturer's overview of current service provision

The description of the current treatment options for people with RRMS appears to be appropriate. The ERG clinical advisors state that the mainstay of current disease modifying treatment (DMT) is interferon beta (IFN-β) and glatiramer acetate (GA) however it is acknowledged that their efficacy is limited. Currently the preferred treatment for active RRMS is natalizumab and another treatment option is fingolimod. The submission outlines the current NICE guidance for treatment with natalizumab and fingolimod. The ERG advisors note that some patients do not receive DMTs initially. This can be due to geographic variation and poor access to specialist services, although the situation is changing now, especially when there is evidence of active MS.

Superseded – see Erratum

## 2.3 Critique of manufacturer's definition of decision problem

### Population

The NICE scope states that the population should have RRMS and the GMS states the population of relevance is those with 'active' RRMS based on the inclusion criteria of the included trials, in which participants had to have had a relapse within the last 12 months. According to the ERGs expert advisors most people with RRMS have 'active' disease, however there are some with very mild or quiescent MS who this would exclude. The ERG advisors suggest that the GMS decision problem is clinically appropriate.

### Intervention

Currently alemtuzumab does not have marketing authorisation, although a positive Committee for Medicinal Products for Human Use (CHMP) opinion was granted in June 2013 and the indication for treatment is anticipated being for RRMS with active disease defined by clinical or imaging features. The manufacturer anticipates alemtuzumab being used in treatment naïve patients and for those who have failed existing therapies. The recommended dose of 12 mg/day via intravenous (IV) infusion in two treatment courses (initial one for five consecutive days, second 12 months later for three consecutive days) appears reasonable. For some patients more than two courses will be required and this is reflected in the GMS (the proportion of patients requiring this is discussed subsequently in the HE section).

### Comparators

The comparators included in the decision problem reflect those in the scope. These are IFN-β, GA, natalizumab and fingolimod. There are restrictions on the use of two comparator interventions (fingolimod and natalizumab), with regard to the specific patient populations covered by their marketing authorisation and NICE guidance (discussed more fully below). The GMS also applies a wider use of these two comparators in the decision problem.

### Outcomes

The outcomes appear to be appropriate to the decision problem. The key outcomes of relevance are relapse rate, severity of relapse, disability, symptoms, freedom from disease activity, mortality, adverse events, health-related quality of life and hospitalisations.

Superseded – See Erratum

**Other relevant factors**

There are two subgroups noted in the decision problem, those with highly active RRMS despite treatment with IFN-β, and those with rapidly evolving RRMS. The NICE scope specified four subgroups, including the two noted in the GMS decision problem, and also treatment naive and previously treated populations. The GMS comments that no subgroup analysis in the cost effectiveness analysis was undertaken and that results were pooled from both treatment naïve and previously treated populations 'to capture more appropriately use across the broad range of the license'. This is discussed in more detail in subsequent sections.

There are no obvious issues related to equity or equality in the decision problem.

# 3 CLINICAL EFFECTIVENESS

## 3.1 Critique of manufacturer's approach to systematic review

**Description of manufacturer's search strategy**

The GMS search strategies are considered overall to be of a reasonable quality, with a few minor inconsistencies. There is a mix of free text and descriptors that have been correctly combined into sets on appropriate databases. The PICO (participants, intervention, comparator, outcomes) breakdown within the search strategy line numbers throughout the GMS was considered to be useful. There is some slight variance in the text and appendices of reporting the dates the searches were undertaken.

The GMS used Pubmed in the clinical search strategy and Medline in all other searches. The GMS did not report the number of return hits per line number in the searches for clinical effectiveness, but did for the other searches. In addition, Embase is not cited for the clinical searches, but is reported in all other searches. The use of the same platform for all the searches would have been a more consistent approach.

The ERG uses Ovid as a search interface and this employs a slightly different syntax; however, the search strategies appear to be appropriate. There were some differences in the terms used to represent the intervention and comparator elements of the clinical searches and the cost related searches. The GMS has elected to use a highly specific RCT filter in the clinical

Superseded - see Erratum

searches, although the results on testing with a more sensitive RCT filter appeared not to produce additional significant results.

The searches were not updated prior to submission and the ERG has therefore updated the clinical and cost effectiveness searches up to end July 2013 (see below for details).

There did not appear to be a systematic search for ongoing studies. The ERG ran searches on UKCRN, Current Controlled Trials, clinicaltrials.gov and WHO ICTRP; results are discussed below.

The GMS does not report a separate search to identify adverse drug reactions. This appears a reasonable approach as the ERG considers that adverse event search filters are of questionable value and that side effects are not always reported in abstracts on bibliographic databases. The text on GMS page 186 indicates that safety data were pooled from the main trials, CAMMS223,[1] CARE-MS I,[2] CARE-MS II[3] and a trial extension (CAMMS03409[4]), and from the manufacturer's safety update reviews.

A reasonable range of grey literature has been searched to identify conference abstracts throughout the GMS and hand searching has also been reported. The quality of life searches use an acceptable filter with restrictions to the UK on one platform but not on the NHSEED database, which appears reasonable. The natural history epidemiology searches and the mortality searches appear appropriate with a range of grey literature searched.

**Statement of the inclusion/exclusion criteria used in the study selection.**

The inclusion and exclusion criteria are clearly stated in Table B6.2.1 of the GMS (page 66). This appears to be relevant to the clinical effectiveness review and the mixed treatment comparison (MTC). As noted above the GMS stated that the focus in the decision problem was on 'active' RRMS, however, in the inclusion criteria this is not stated as such, text states 'adult patients with RRMS'.

Single or double blind randomised controlled trials (RCTs) and open label extensions of RCTs were eligible for inclusion. No limits for inclusion were placed on eligibility relating to study quality and setting was not used as an inclusion criterion, but this does not appear to be a relevant factor.

The NICE scope for this appraisal requires an assessment of alemtuzumab compared with IFN-β, GA, natalizumab (for treatment naïve or previously treated rapidly evolving severe [RES] RRMS) and fingolimod (for those with previously treated highly-active RRMS). The GMS has each of the stated treatments in their inclusion criteria, also eligible were BG-12, daclizumab, laquinimod, mitozantrone, rituximab, and teriflunomide. These were outside of the scope of the assessment however were included in the MTC (see below for further discussion of their relevance).

Three RCTs were included (further details are provided below), of which the population were treatment naïve in two and previously treated in one. For the comparison with IFN-β all three trials are appropriate. For the comparison with GA there was no direct evidence, however, through an indirect comparison (see discussion of MTC) the three RCTs were appropriate. For the comparison with natalizumab the two RCTs with previously untreated populations are appropriate to be compared through indirect comparison. There is some uncertainty over the relevance of the RCT of previously treated participants to the comparison with natalizumab (through indirect comparison) as the population for this comparison should be those with RES RRMS. This trial was reported (GMS page 102, 125) to have some RES RRMS participants, and subgroup analyses were presented, however, the complete population was used for the evidence for this comparison in the MTC. This is discussed in more detail in subsequent sections. For the comparison with fingolimod (via indirect comparison), the RCT undertaken on participants who had been previously treated may be relevant to the decision problem, however, there is some uncertainty as to how many participants had highly active RRMS. The complete population was used for the evidence for this comparison in the MTC. These factors are discussed in more detail in subsequent sections.

The ERG clinical advisors state that the populations in the trials are reasonably typical of those likely to receive DMTs in the UK. The ERG note that in Table B6.3.4 (GMS page 87-9) of the GMS the CAMMS223[1] trial had approximately 10% of participants with one or zero relapse in the past two years when the inclusion criteria states this should be at least two. Also, there were participants in the CARE-MS I[2] and II[3] trials that do not appear to meet the inclusion criteria based on the number of relapses in the previous two years. Clarification received from the manufacturer confirms that for a small number of patients there was a discrepancy between patient recruitment and the inclusion criteria for CAMMS223,[1] CARE-MS I[2] and CARE-MS II[3] in

Superseded – see Erratum

relation to meeting the required inclusion criteria for frequency of MS episodes prior to trial entry. According to the clarification, patients were recruited on the basis of having two episodes in the previous two years and only on review of their notes did it became clear that in some cases these episodes did not fulfil the protocol definition of a relapse (ClarificationA4 page 6). The clarification response also describes an adjusted analysis which accounts for some of this deviance from the inclusion criteria. This is discussed in more detail below.

In addition, Table B6.3.4 of the GMS (page 87-9) provides details of the proportion of participants with different EDSS scores at baseline. It was difficult to ascertain from this table whether the participants met the respective inclusion criteria of the three RCTs because of the categories of EDSS used. Clarification provided from the manufacturer has shown that a small proportion of individuals had EDSS scores at baseline that were outside of the stated inclusion criteria (see Table 1). The clarification states that the EDSS scores at screening could differ from baseline, but are unable to provide the screening EDSS scores (Clarification A5 page 7).

The eligible outcomes appear to be reasonable and appropriate for the assessment, and although limited description is provided in Table B6.2.1, more detail is provided in the Appendices (Section 10.2) of the GMS.

A flow chart (GMS page 68) with the numbers included and excluded at each stage, meeting the criteria of the PRISMA statement was presented and appears to be correct where it can be cross checked.

A list of excluded studies was not presented so reviewers are unable to check whether any studies were excluded inappropriately.

The ERG is not aware of any potential bias in the selection of studies.

**Identified studies**

Three RCTs were identified in the submission that relate to alemtuzumab:

1) CAMMS223 (Coles and colleagues 2008[1]) is an RCT comparing subcutaneous (SC) IFNβ-1a (44µg) with IV alemtuzumab 12 mg/day and IV alemtuzumab 24mg/day. The population were treatment naive.

2) CARE MS-I (Cohen and colleagues 2012[2]) is an RCT comparing SC IFNβ-1a (44µg) with IV alemtuzumab 12 mg/day. The population were treatment naive.
3) CARE MS-II (Coles and colleagues 2012[3]) is an RCT comparing IFNβ-1a (44µg) with IV alemtuzumab 12 mg/day and IV alemtuzumab 24mg/day. The population had been previously treated with DMTs (see below for more discussion of this).

Therefore there is limited direct evidence comparing alemtuzumab with other DMTs relevant to the decision problem. In Section 6.7, the GMS included an additional 27 RCTs of comparator treatments in a MTC. These consisted of:

- 1 RCT of daclizumab (x2 doses) X placebo;
- 5 RCTs of fingolimod (x2 doses) X placebo (4) or IFNβ-1a (1);
- 3 RCTs of BG-12 X placebo (1 also comparison of GA);
- 2 dose ranging RCTs of GA (2 doses, no placebo);
- 2 RCTs IFNβ-1b X GA (1 with 2 doses of IFNβ-1b);
- 2 RCTs IFNβ-1a X GA (1 also with a comparison of IFNβ-1b and GA);
- 2 RCTs comparing 2 or 3 different types of IFNβ respectively;
- 1 RCT comparing laquinimod with placebo or IFNβ-1a;
- 3 RCTs of laquinimod (1 x2 doses) X placebo;
- 1 RCT natalizumab X placebo;
- 1 RCT rituximab X placebo;
- 3 RCTs teriflunomide (x2 doses) X placebo;
- 1 RCT teriflunomide (x2 doses) X IFNβ-1a.

No RCTs have been included that do not appear to meet the inclusion criteria. RCTs of BG-12, laquinimod, rituximab and teriflunomide were not relevant to the decision problem. In the case of teriflunomide a connection was made through this intervention to compare alemtuzumab with natalizumab in the MTC. The other treatments were not required for any indirect comparison. For a full discussion of the MTC, see below.

Summary details of the three RCTs were provided in the GMS. Table B6.2.3 (GMS page 71) summarises the design, interventions and dosage information. Table B6.3.1 (GMS page 74) summarises key detail on the objectives, study design, location, recruitment, study duration, methods, intervention details and outcomes, and the duration of follow-up. Flow-charts with

Superseded – see Erratum

patient numbers are reported for each of the three RCTs (GMS pages 104-106). A summary of the statistical analyses are provided in Table B6.3.7 (GMS page 96) and subgroup analyses are described on pages 96-102.

The ERG has checked the information provided in the GMS with the trial publications and clinical study reports (CSR) where available. There are a few issues of note from the three alemtuzumab RCTs:

Alemtuzumab treatment in trial CAMMS223[1] was suspended after three cases of immune thrombocytopenic purpura (ITP), including 1 fatality. The trial was later resumed (suspended September 2005, recommenced April 2008). Treatment with IFNβ-1a continued and all participants, regardless of treatment arm, continued with assessments of efficacy and safety. At the time of the suspension, two eligible participants had not received the second cycle of alemtuzumab at month 12, and 155 participants were precluded from receiving the third dose at month 24 (not reported in the GMS, from main publication[1]). Figure B6.3.1 (GMS page 81) shows the trajectory of participants, however there is limited information about the assessments made to recommence treatment. It is also unclear what the mean length of follow-up for participants was at the point of the suspension or what proportion of participants were included in the three year efficacy and safety analysis. When treatment with alemtuzumab was resumed the trial was in an extension phase[4] and participants who had previously been in the two alemtuzumab trial arms [redacted]. Any analysis after the recommencement was therefore not directly related to the original randomisation schedule. Clarification received from the manufacturer explained that after the trial suspension, an independent unblinded Data Safety Monitoring Board (DSMB) implemented a risk minimisation action plan. Once the DSMB were sufficiently reassured by the measures put in place the dosing suspension was lifted and patients could re-enter the study if they did not have any of the disqualifying criteria (Clarification A2 page 2, including Table B5 omitted from the GMS). Data analysis of clinical efficacy was carried out on a yearly basis, but no efficacy analysis relating specifically to the point of alemtuzumab dosing suspension was undertaken (Clarification A2 page 5). No participants who did not receive second or third doses were given retreatment prior to the 36 month evaluation, Clarification A2 page 4).

A range of tertiary end points were listed for these studies in the GMS (Table B6.3.1 page 78). In the CAMMS223[1] trial these were not reported as such in the trial publication. However, these were later identified in the protocol sent to the ERG as part of the clarification stage. For CARE MS-I[2] some pre-specified tertiary endpoints were reported in the publication and these concur with those presented in the GMS.

The baseline characteristics for CARE MS-I[2] and CARE MS-II[3] are based on the per-protocol treated populations rather than the randomised populations. This was also the case in the trial publications. Eighteen and 42 participants in the two trials respectively were therefore excluded from the baseline assessments.

Baseline characteristics appear to be similar between the two intervention groups in the three RCTs. Some small differences can be seen in CAMM223[1] where the mean ages were similar but age ranges varied from 18-60 years in the IFN-β group and 18-49 years in the alemtuzumab 12 mg group. In addition the history of relapse had similar median rates, but the ranges varied (from 0.2-6.3 years in the IFN-β group and 0.1-3.5 years in the alemtuzumab 12 mg group). Finally, relapse in the previous two years of three or more was 27% in the IFN-β group and 41% in the alemtuzumab 12 mg group.[1]

There are some differences in participants between the three RCTs owing to differences in the inclusion criteria. The GMS (page 85) points out that all participants in the CAREMS-II[3] trial had received previous DMTs (discussed more below) whereas the populations in the CAMMS223[1] and CARE MS-I[2] had not. Also participants in CARE MS-II[3] had a greater time since first relapse compared to the populations in the other two RCTs. Participants in CAMMS223[1] and CARE MS-I[2] had early active RRMS (defined as Expanded Disability Status Scale [EDSS] scores ≤3) and duration of disease of either ≤3 years or ≤5 years respectively, with at least two clinical episodes in the previous two years (additional specific criteria in the CAMMS223[1] trial were ≥1 clinical episode in the previous year and ≥1 gadolinium enhancing lesion on cerebral MRI). In the CARE MS-II[3] trial participants could have had MS for up to 10 years and could have a EDSS score up to five.

As a result of these differences in inclusion criteria, the following differences were observed in the baseline characteristics of the three trials:

- The mean time since first episode ranged between [redacted] in CAMMS223;[1] 2.0-2.1 years in CARE MS-I;[2] and 4.3 – 4.7 years in CARE MS-II.[3]
- Similarly, the median time since the *first* relapse ranged between 1.2-1.4 years in CAMMS223;[1] [redacted] years in CARE MS-I;[2] and [redacted] years in CAREMS-II.[3]
- Participants in the CAMMS223[1] trial had a single relapse in the previous two years of range 5.4 - 11.8% [redacted].
- Mean EDSS scores ranged from 1.9 - 2.0 in the CAMMS223[1] trial; [redacted]For more precise breakdown of EDSS scores see GMS Table B6.3.4 (page 88). [redacted]The mean T2 lesion volume ranged from [redacted] in CAMMS223;[1] 7.3 - 7.4cm$^3$ in CARE MS-I;[2] and 9.04 - 9.94 cm$^3$ in CAREMS-II.[3]

All baseline characteristics have been checked with the trials where data were reported. Some data in CARE MS-I[2] and CARE MS-II[3] are marked CIC in the GMS but are available in the trial publication (median time since first relapse, mean EDSS scores). As noted above, the baseline characteristics for CARE MS-I[2] and CARE MS-II[3] are reported in the trials and the GMS for those who received at least one dose of the study medications (i.e the per protocol population rather than the randomised population).

In the CARE-MS II trial[3] participants had been previously treated with DMTs. The GMS reports details of the type of DMT participants in each arm had received in Table B6.3.4 (p87-9). From these data it can be seen that the majority of participants received either SC IFNβ-1a; IM IFNβ-1a; SC IFNβ-1a (22µg or 44µg); IFNβ-1b or GA. The comparator in the CARE-MS II[3] was SC IFNβ-1a (44µg) and therefore some participants in this trial had already been unresponsive to treatment with IFNβ-1a (44µg), and many participants had been unresponsive to other types of interferon treatment. This should be considered when interpreting the results of the comparison of alemtuzumab with IFNβ-1a (44µg).

All other baseline characteristics appear to be similar between the three RCTs. Table 1 presents the key baseline characteristics for the three RCTs.

**Table 1: Main characteristics of participants in the three included RCTs**

| | **CAMMS223[1]** | | **CARE-MS I[2]** | | **CARE-MS II[3]** | |
|---|---|---|---|---|---|---|
| | SC IFNβ-1a (44µg) | Alemtuzumab (12mg)[a] | SC IFNβ-1a (44µg) | Alemtuzumab (12mg) | SC IFNβ-1a (44µg) | Alemtuzumab (12mg)[a] |
| **Number** | 111 | 112 | 187 | 376 | 202 | 426 |
| **Mean age (SD) years** | 32.8 (8.8) | 31.9 (8.0) | 33.2 (8.5) | 33.0 (8.0) | 35.8 (8.77) | 34.8 (8.36) |
| **Female (%)** | 64.0 | 64.3 | 65.2 | 64.6 | 64.9 | 66.0 |
| **White (%)** | 90.1 | 91.1 | 96.3 | 93.6 | 92.6 | 90.4 |
| Relapse in previous 2 years (%): | | | | | | |
| 0 | 0 | 2 (1.8) | [redacted] | [redacted] | [redacted] | [redacted] |
| 1 | 8 (7.2) | 6 (5.4) | [redacted] | [redacted] | [redacted] | [redacted] |
| 2 | 73 (65.8) | 58 (51.8) | [redacted] | [redacted] | [redacted] | [redacted] |
| ≥3 | 30 (27.0) | 46 (41.1) | [redacted] | [redacted] | [redacted] | [redacted] |
| **Mean (SD) relapses** | - | - | [redacted] | [redacted] | [redacted] | [redacted] |
| **Mean duration of previous MS drug use in months (SD)** | - | - | - | - | 36 (23.7) | 35 (25.0) |
| **Mean EDSS score** | 1.9 (0.83) | 1.9 (0.74) | [redacted] | [redacted] | [redacted] | [redacted] |
| 0 | 8 (7.2) | 4 (3.6) | [redacted] | [redacted] | [redacted] | [redacted] |
| >0-1.5 | 37 (33.3) | 40 (35.7) | [redacted] | [redacted] | [redacted] | [redacted] |
| >1.5-2.0 | 28 (25.2) | 30 (26.8) | [redacted] | [redacted] | [redacted] | [redacted] |
| >2.0-3.0[b] | | 38 (33.9) | 65 (35) | 140 (37) | 48 (24) | 112 (26) |
| >3.0-4.0[b] | - | 0 | 3 (2) | 8 (2) | 50 (25) | 98 (23) |
| >4.0-5.0[b] | - | - | - | - | 19 (9) | 42 (10) |
| >5.0-6.0[b] | - | - | - | - | 2(1) | [redacted] |
| >6.0-7.0[b] | - | - | - | - | 0 | [redacted] |

[a] Also had a third treatment arm with alemtuzumab 24 mg not reported here as not the anticipated licensed dose. [b] Based on Clarifications (A5.1 page 8).

Some information was not available to be cross-checked with the trial publication.

All three RCTs[1;3;4] had follow-up of at least two years, which is in line with European Medicines Agency (EMA) recommendations that trial duration should be at least two years to allow enough time to assess relapse.[5]

All of the included studies were sponsored by the manufacturer of alemtuzumab and the manufacturer of IFNβ-1a.

In addition to the three RCTs, data from an extension to CAMMS223[4] and an extension to all three alemtuzumab trials[6] were included in the GMS. Two other non-RCTs were also reported.

[redacted] The other was a prospective study of data from patients treated with alemtuzumab at five UK centres who were not included in the three pivotal RCTs.

[redacted] The other study did not report specific inclusion criteria, but baseline characteristics show 43.1% had been treated with DMTs. It is therefore likely that the participants in these studies were a mixture of untreated and previously treated.

It is unclear how these two non-RCTs were identified without a systematic review (GMS page 180) and without a search for non-RCT evidence (GMS Appendix 6, page 432). It is unclear whether any other non-RCTs would be relevant to the decision problem or whether these two studies were more relevant than other non-RCTs, [redacted] The GMS stated that a systematic review was not undertaken due to the weight of evidence from the RCTs and it is unclear why this particular data was included. Due to these reasons, the ERG did not assess the non-RCTs or their data, but note that in these trials approximately 30[redacted] (GMS page 181) of participants received three or more courses of alemtuzumab as this may be relevant to the economic model.

The date of the last search for clinical effectiveness data was November 2012 (see Section3.1) and the ERG have updated the GMS searches until end of July 2013. 246 references were identified. No formal screening procedures were applied however, the ERG have identified 5 studies that may be relevant that were not identified by the manufacturer (cross checked with alemtuzumab and MTC interventions, Table B6.7.2 and 3), although the full list of excluded studies was not provided to check these. While some of the interventions may not be relevant to the scope, they may have been eligible for the MTC. The ERG has been unable to assess these RCTs further.

**Table 2 Potential additional studies identified by the ERG**

| Author, year | Participants | Intervention | Comparator |
|---|---|---|---|
| Benedetti *et al.*, 2012[7] | RRMS | Azathioprine | IFN |
| DeStefano *et al.*, 2012[8] | RRMS | IFN beta-1a | Placebo |
| Hauser *et al.*, 2012[9] | RRMS | Ocrelizumab, or IFN | Placebo |

| | | | |
|---|---|---|---|
| Khan *et al.*, 2013[10] | RRMS | GA | Placebo |
| Massacesi *et al.*, 2012[11] | RRMS | Azathioprine | IFN |

Ongoing studies searches by the ERG have identified no ongoing studies of relevance.

**Description and critique of the approach to validity assessment**

The GMS applied the NICE quality assessment questions to the three included RCTs. The ERG have similarly applied the questions to the trials and have tabulated the ERG responses alongside those of the GMS in Table 3.

The GMS reports that allocation was by interactive voice response system (IVRS) for all three trials, with concealment of treatment allocation assessed as adequate by the ERG. Although data to check this for CAMSS223[1] were not available in the published trials this was identified in the Clinical Study Report (CSR) sent at the clarification stage.

The submission reports that there was no blinding of care providers, participants or outcome assessors in the three RCTs. The ERG agrees that it would not be possible to blind care providers or participants, but that outcome assessors could be blinded. In the three RCTs some attempt has been made to blind the outcome assessors, although this appears not to be complete for all outcomes and has therefore rated this as partial. This is of particular concern for the reporting of subjective outcome measures such as the EDSS (see below for more discussion of this outcome) and EMA recommendations that the identification of relapses should be blinded to therapy.[5]

With the information available to the ERG based on the full CSRs, there seems to be a low risk of reporting bias.

The ERG have assessed the two CARE MS RCTs[2;3] as not applying an ITT analysis, as only those receiving at least one dose of medication were analysed and a proportion of individuals from each group did not receive study medication. Clarification received from the manufacturer to a question about the participants in the included studies (above) stated that for the primary efficacy analyses all three trials used the 'full analysis set' (all patients who were randomised and received any amount of study drug). A 'per-protocol set' (for criteria see Clarification A4 page 6-7) was used for analysis of the co-primary endpoints if it was <90% of the full set, as

was the case in CAMMS223[1] [redacted] and Care-MS II[2] [redacted]. This was not the case in CARE-MS I[3] [redacted] Therefore analyses of the co-primary endpoints (and possibly other endpoints) are based on an incomplete data set. In addition, as noted above, in some cases the trials included participants that did not fully meet the inclusion criteria.

Overall there is a low risk of selection bias in these trials, but an uncertain risk of detection bias which should be taken into account when interpreting the results.

**Table 3: Manufacturer and ERG assessment of trial quality**

| | | CAMMS223[1] | CARE MS I[2] | CARE MS II[3] |
|---|---|---|---|---|
| 1. Was randomisation carried out appropriately? | MS: | Yes | Yes | Yes |
| | ERG: | Yes | Yes | Yes |
| Comment: CAMMS223 used minimisation with a random component so likely to be low risk of bias, randomisation was also stratified by site<br>CARE MS I and II stratified by centre but used a centralisation schedule. | | | | |
| 2. Was concealment of treatment allocation adequate? | MS: | See comments | See comments | See comment |
| | ERG: | Yes | Yes | Yes |
| Comment: The manufacturer does not give a direct answer of yes/no/unclear. They report that for these three trials allocation to treatment was randomly assigned by IVRS. Treatment group was not concealed from patients and clinicians as study drugs had distinctive adverse effects that precluded masking assignment. Clinical data integrity was however secured by stringent clinical and MRI rater masking<br>CAMMS223: trial publication does not specify how allocation was concealed but CSR reports IVRS use.. | | | | |
| 3. Were groups similar at outset in terms of prognostic factors? | MS: | Yes | Yes | Yes |
| | ERG: | Yes | Yes | Yes |
| Comment: some small differences between groups for CAMMS223 | | | | |
| 4. Were care providers, participants and outcome assessors blind to treatment allocation? | MS: | No | No | No |
| | ERG: | Partial | Partial | Partial |
| Comment: CAMMS223 attempted to blind the outcome assessors for one outcome but not reported for all. CARE MS I outcome assessment was masked, although some question over whether this was always adequate.<br>CARE MS II outcome assessment was masked, however, unclear whether for all outcomes and small proportion reported to not have been adequately masked. | | | | |
| 5. Were there any unexpected imbalances in drop-outs between groups? | MS: | Yes | Yes | Yes |
| | ERG: | Yes | Yes | Yes |
| Comment: | | | | |
| 6. Is there any evidence that authors measured more outcomes than reported? | MS: | No | No | No |
| | ERG: | No | No | No |
| Comment: Assessment based on the full CSRs. | | | | |
| 7. Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? | MS: | See comments | See comments | See comments |
| | ERG: | Yes | No | No |
| Comment: The manufacturer does not give a direct answer of yes/no/unclear. They report that for these three trials the Full Analysis (FA) Set (all randomised patients who had a diagnosis of MS) was the primary population and missing data was accounted for.<br>CAMS223 – one participant was excluded from the analysis as incorrect diagnosis, however, unlikely to bias results. Was included in the safety analysis. | | | | |

| |
|---|
| CARE MS I and CARE MS II – state ITT but analysed only those who had received at least one dose of study medication which suggests a modified ITT analysis. |

**Description and critique of manufacturer's outcome selection**

The scope stated outcomes of relapse rate, severity of relapse, disability (e.g. EDSS), symptoms of MS, freedom from disease activity, mortality, adverse events and health-related QoL. These were all reported as outcomes in the decision problem of the GMS. The key outcomes reported in the clinical effectiveness section of the GMS were relapse rate, annualised relapse rate (ARR) and sustained accumulation of disability (SAD). Primary outcomes in the trials were time to 6 months SAD and rate of relapse for CARE MS I[2] and II,[3] and time to SAD and rate of relapse for CAMMS223.[1]

The EMA draft guidelines for MS 2012[5] state that ARR is an acceptable outcome. Also discussed in the guidelines is that the relapse can be difficult to identify because patients can suffer from pseudo-exacerbations. The report therefore suggests that the definition of relapse should include occurrence, time of start and end, a minimum duration, a maximum time between two symptoms to qualify as a single relapse, and severity. The GMS provides details of relapse definitions used in the trials which appear to cover most of these requirements.

Additional outcomes reported in the GMS were hospitalisation as a result of relapses.

The GMS focuses on the outcomes from the scope. A range of secondary and tertiary outcomes were also reported to have been analysed in the trials, see Table B6.3.1 (GMS pages 74-9) for details. No data were presented in the GMS for these outcomes. One of these, EQ-5D, was reported to be published in an abstract and may have had relevance to the economic evaluation, but no data were reported.

The SF-36 and Functional Assessment of Multiple Sclerosis (FAMS) scales were used to assess QoL, both of which are validated tools.

There are known limitations with the EDSS, however, there is no suitable alternative measure and the EMA suggest it should be used for comparison between studies.[5] The mean change in score from the baseline is not recommended to assess efficacy, however, a predefined level of change to indicate treatment failure or progression should be used. For example, the achievement of a specified degree of disability or sustained worsening (the EMA suggest 1 point

when EDSS scores ≤ 5.5; 0.5 points if baseline score is > 5.5[5]). The GMS reported the SAD based on the EDSS and this was defined as an increase of at least 1.5 points if the baseline EDSS was zero, and an increase of at least one point for patients with a baseline score of one or more (GMS page 91).

**Description and critique of the manufacturer's approach to trial statistics**

The GMS focuses on results for the alemtuzumab 12 mg dose rather than the 24 mg dose from the two trials that included a third treatment arm (CAMMS223[1] and CARE MS II,[3]) because the 12 mg dose is anticipated to be used for standard treatment and was submitted to the CHMP for consideration of use in practice. The ERG has similarly focused on the alemtuzumab 12 mg dose results (see below for study results).

There are some variations in the reported outcomes and details of results for the three RCTs. All three RCTs are described as using ITT methods of analysis. However, as shown in Table 3 above in two RCTs[2;3] the population assessed appeared to be from a modified ITT group including all randomised patients that received at least one dose of study medication.

Expanded trial results reporting primary and secondary outcomes were presented in GMS Additional Appendix 3.

The GMS (B6.3.7 page 98, 99) states that in the CARE MS-I[2] and -II[3] RCTs the assessment of time-to-event endpoints patients were censored at their last visit if the event had not occurred but no further detail is provided.

Adverse events of alemtuzumab 12 mg were presented (GMS Section 6.9, page 185) based on the pooled data from the three key RCTs (CAMMS223,[1] CARE-MS I,[2] CARE-MS II,[3]) and the long-term extension study (CAMMS03409[6]). The individual study data were presented in Additional Appendix 9. The GMS states that the safety and adverse events were based on an 8.5 year interim analysis from the extension Study (CAMMS03409,[6] page 27). Safety data were presented through to 31.12.2011 with additional data from the Safety Update Report (SUR), which provides updated data from CAMMS03409[6] through to 26.11.2012. Very limited data on the safety of IFN-β from these studies are presented in the main GMS.

Four subgroups were defined in the NICE scope (treatment experienced RRMS; treatment naive RRMS; HA RRMS; RES RRMS). Subgroup analyses were undertaken for a population with RES RRMS from the CAMMS223[1] and CARE-MS II[3] trials. This was defined as ≥2 relapses in the year prior to treatment; at least 1 gadolinium-enhancing lesion at baseline. This differs slightly to the definition of RES RRMS in the licensed indication and NICE guidance for natalizumab (NICE 2007): ≥ 2 disabling relapses in one year with ≥1 gadolinium-enhancing lesions on brain MRI *or a significant increase in T2 lesion load as compared to a previous recent MRI.* The subgroup analyses were pre-planned for some outcomes but post-hoc for others as follows:

- ARR (pre-planned)
- SAD (pre-planned)
- Sustained reduction in disability, clinical disease activity free and MRI activity free subgroup analysis (post-hoc, 60 months after randomisation, GMS page 96).

No discussion of the interpretation or clinical meaningfulness of the results was presented. There is little discussion or justification in the GMS of any clinically important differences between the RCTs.

**Description and critique of the manufacturer's approach to the evidence synthesis**

Apart from the long-term data, the narrative review of the GMS only presents a summary sentence of positive results per RCT for each of ARRs, SAD, EDSS, MRI imaging outcomes, Freedom from disease and QoL (GMS pages 112-113) with the remaining results presented in tabulated form (GMS B6.5.1, pages 114-9). The GMS Table B6.5.1 contained an error, in that 6–month SAD at 3 years was reported twice for CAMMS223[1] with different data. The second 6-month SAD at 3 years should have read 3-month SAD at 3 years (Clarification A6 page 8-9). Otherwise, the data presented appears to fully reflect the data in the RCTs, a full check of the data from the extension studies or other non-RCTs was not made.

The GMS undertook a meta-analysis of the three included RCTs (CAMMS223,[1] CARE-MS I[2] and II[3]). Outcomes were combined for ARRs, proportion of patients relapse free, 3 and 6 months SAD, all-cause discontinuations and discontinuations due to adverse events (DAEs). In the narrative the comparator from the meta-analysis was stated to be placebo (page 131) which has been clarified by the manufacturer as being a typographical error and should read SC IFNβ-1a (Clarification response A8, page 9).

The ERG considers that it is not appropriate to combine the three RCTs because of differences in the populations between the three studies. Two of the RCTs included treatment-naïve participants (CAMMS223[1] onset of MS within past 3 years; CARE-MS I[2] MS onset within 5 years) and CARE-MS II[3] included recently relapsed treatment-experienced patients, (MS onset within past 10 years and treated with beta interferon or GA). In addition, there were differences in the inclusion criteria for EDSS scores with corresponding differences in the included populations. This differed for CARE-MS II[3] (EDSS 0.0 to 5.0) compared to CAMMS223[1] and CARE-MS I[2] (both EDSS 0.0 to 3.0), see also Table 1. The MS episode history also differed between studies. In the CAMMS223[1] trial the criteria were ≥2 relapses in the 2 years prior to the study, which was different to the CARE-MS I[2] and CARE-MS II[3] trials which required ≥2 relapses in the 2 years prior to the study plus at least 1 episode in year prior, with objective neurological signs. The GMS noted that CARE-MS II[3] had treatment-experienced patients with a greater time since first relapse compared to the other 2 RCTs[1;2] but this was not explored in any sensitivity analyses (see below).

A statistical assessment of heterogeneity is reported for each outcome (all data were marked CIC). Demographic patient details of the RCTs were assessed manually. The GMS states that generally, very few signs of heterogeneity were found using informal classical meta-analyses of direct and one-bridge-indirect evidence. The ERG requested clarification over how the meta-analysis included one-bridge indirect evidence. The clarification describes this in the context of the MTC and therefore the ERG assume that this statement was related to the MTC rather than the meta-analysis (see below for discussion of MTC; Clarification A9 page 9). Discontinuations due to adverse events were the exception, which according to the GMS lacked power to detect any differences.

The GMS reports using a random-effects model and employed an empirical Bayes estimator of the random-effects variance. No justification for the choice of model was provided however this was provided on clarification. The ERG note that fixed effects meta-analyses were not presented in the GMS as per the NICE submission template (Clarifications A7 page 9). Very little description of the methods was provided overall.

The combined results are reported as relative risks (RR), see ERG report page XX. The GMS presented RR and 95% confidence intervals (CI) stating that the results were statistically

Superseded – see Erratum

significant (in that the CIs do not include 1) [redacted]. No forest plots were presented for the direct meta-analysis, the GMS refers to the network meta-analysis (section 6.7) for any forest plots.

No subgroups were analysed in the meta-analysis and no sensitivity analysis of the different populations (for example, treatment naïve /previously treated) was presented.

The GMS included a MTC consisting of 30 clinical trials with patient recruitment after 2000 and ≥80% RRMS patient population (see below for specific discussion of the MTC). Section 6.7.1 stated that sensitivity analyses were run on datasets without these restrictions, (on the dataset ≥2000 recruitment year and 100% RRMS patients, GMS page 131) referring to Additional Appendix 4 for further details.

Most of the methods were described. The justification for excluding studies with patient recruitment before 2000 due to decreased relapse rates over time was that this was based on advice from clinical experts and coincided with the widespread introduction of the McDonald diagnostic criteria in clinical trials (page 134). An Advisory Board reviewed the distribution of the proportion of RRMS patients in the included trials and an analysis of the trials found that the vast majority appeared to cluster into two groups: ≥80% RRMS patients or considerably <80% RRMS patients. As such, 80% was selected as the cut-off for analysis. The ERG clinical advisors suggest this is reasonable.

The GMS does not explicitly state the rationale for conducting MTCs when reporting the methods of analysis. However they acknowledge, when discussing the relevance of the evidence base to the decision problem, that a lack of head-to-head comparisons for included treatment regimens required some form of indirect comparison. It would also appear that the structure of the economic model required estimates of the effectiveness of alemtuzumab relative to placebo (treated as best supportive care) to be derived, rather than the trial-based comparisons with SC IFNβ1-a.

A summary of the MTC methodology (inclusion/ exclusion criteria [GMS section 6.7.1], evidence networks [GMS section 6.7.4] outcomes and methods [GMS section 6.7.6]) is presented in the GMS and further details including data inputs and WinBUGS code in Appendices.

The ERG appraised the methodological quality of the MTC (Table 4).

**Table 4: ERG appraisal of MTC approach**

| Appraisal criteria | Criteria met (YES / NO / UNCLEAR / NOT APPLICABLE) |
|---|---|
| **A. CONCEPTUAL BASIS** | |
| 1. Is a justification given for conducting an MTC? | There is no rationale/ justification stated for conducting MTC in section 6.7 of GMS reporting the methods for the indirect and MTCs. The rationale for conducting MTC is acknowledged explicitly in paragraph 2 on page 218 of GMS – "evidence is presented from a MTC as head-to-head data are absent". |
| **B. SYSTEMATIC PROCESSES** | |
| 2. Is a comprehensive and transparent search strategy reported? | No search specific to the MTC was conducted. However, the search for the systematic review of clinical effectiveness included a range of comparator DMTs which were relevant to the MTC. Searches were completed approximately 7 months prior to submission and the ERG have updated searches (see above) |
| 3. Are inclusion / exclusion criteria adequately reported? | Yes (GMS pages 134-135) |
| 4. Is the number of included /excluded studies from the MTC reported, with reasons for exclusions? | The number of included/ excluded studies is reported in a flow chart (GMS page 136). However, it is unclear in how the final list of 30 included trials was arrived at. A trial exclusion list was not provided. Of the subsequently excluded MTC trials, a list was provided with clarifications (Clarification page 15-16). This shows that 19 trials were excluded as recruitment occurred prior to 2000 and/or <80% had RRMS. In addition, 3 trials were excluded because the dosing was outside the product indication. |
| 5. Is a visual representation of the data networks provided? | Yes |
| 6. Are the data from included studies extracted and tabulated? | Yes – data inputs for MTC reported in Appendix 6 |
| 7. Is the quality of the included studies assessed? | Yes. The assessment is tabulated over 14 pages, but no overall synthesis is presented.<br><br>In 3 trials the randomisation procedure is reported to not have been carried out appropriately; in 5 trials it is unclear if the concealment of treatment allocation was adequate; 6 trials had differences in baseline characteristics between treatment arms (plus 1 trial in which it was unclear); not all trials had double-blinding and in some it was unclear if assessors were blinded to treatment for each outcome; 5 trials had unexpected imbalances in drop-out between treatment groups and in a further 7 trials it was judged as unclear; 4 trials appeared to have measured more outcomes than reported; for 5 trials it was unclear if ITT analysis was used and for a further 6 trials it was judged to be unclear if the appropriate methods were used. It is unclear how the assessment was carried out (i.e. 2 independent reviewers). |

| | |
|---|---|
| | The ERG is unable to identify any potential biases as a result of studies being excluded from the MTC. |
| **C. STATISTICAL ANALYSIS** | |
| 8. Are the statistical procedures adequately described and executed? | Statistical procedures are adequately described. The GMS reports that adequate tests were undertaken to assess convergence, though no detail of the results of these tests were provided. The WinBUGS models used for the MTCs do not appear to take account of the inclusion of multi-arm trials.<br><br>The networks appear to include (a) irrelevant comparators that are unconnected (other than to placebo) (b) irrelevant (to the decision problem/ scope) dosages of relevant comparators. There are inconsistencies in inclusion of irrelevant dosages of relevant comparators – with no rationale offered. |
| 9. Is there a sufficient discussion of heterogeneity? | Partial.<br>Clinical heterogeneity (especially changes in diagnostic criteria) and results of heterogeneity tests (and possible sources of heterogeneity not controlled for in the analysis) are discussed. There is no discussion of the appropriateness of pooling treatment-naïve and treatment-experienced participants. |
| 10. Is the type of model used (i.e. fixed or random effects) reported and justified? | Yes. A random effects model is used in all analyses. This is stated but not discussed at any point in the GMS (a justification, based on a recommendation from the Cochrane Collaboration to adopt the more conservative option, of random effect meta analysis was provided in the clarifications [A7 page 9]). |
| 11. Was sensitivity analysis conducted? | Yes (GMS page 157). Sensitivity to decision to exclude trials from before year 2000 and those including non-RRMS patients |
| 12. Is any of the programming code used in the statistical programme provided (for potential verification?) | Yes<br>Analyses have been re-run for 3-month SAD using manufacturer's original code and a modification by ERG to take account of the inclusion of multi-arm trials. |
| **D. PRESENTATION AND INTERPRETATION OF THE EVIDENCE** | |
| 13. Is there a tabulation/ illustration of results for each intervention and for each outcome? | Yes<br>Note that the tabulations and forest plots only present part of the results of the MTC (i.e. tabulations only for comparators in the scope of the decision problem/ economic model and for selected comparators in forest plots). A much wider number of comparisons were analysed in the MTC. |
| 14. Is there a narrative commentary on the results? | The GMS presented no narrative alongside the results of the MTC analyses other than to asterisk "not statistically significant" results (presumably on basis of bounds of 95% credible interval).<br>Commentary on the results of the MTCs is limited to a summary which appears in section 6.10 (interpretation of clinical evidence) of the GMS, comprising two bullet points (page 210) for the base case, and an overview of the sub-group analyses on the same page. |
| 15. Does the discussion of the results reflect the data presented? | The summary presented in section 6.10 broadly reflects the evidence presented in the tables in the relevant earlier sections. |

| | |
|---|---|
| 16. Have the authors commented on how their results compare with other published studies (e.g. MTCs), and offer any explanation for discrepancies? | There are no comparisons to other published meta analyses/ MTCs presented in the GMS. |
| 17. Have the authors discussed whether or not there are any differences in effects between the direct and indirect evidence? | Direct and indirect evidence are discussed separately. The forest plots in section 6.7.7.1 included pooled direct results (from non-network meta-analysis) and the MTC results. The GMS does not include any discussion of consistency. There do not appear to be major discrepancies between the (limited) direct evidence presented (alemtuzumab versus SC IFNβ-1a only) and estimates from the MTC for the analysis of post-2000 trials. There are inconsistencies between direct evidence (excluded from the base case MTC) for SC IFNβ-1a versus placebo and IFNβ-1b 250 µg versus placebo compared with the results derived from the MTC and used in the base case. |

Most of the assessment criteria were met indicating a reasonable approach to the MTC. However there were instances where the criteria were not met or were only partially met. For example, it is not clear how the final list of included trials was arrived, or how robust the results of the analyses may be to variations in trial populations (treatment-naïve versus treatment-experienced and other baseline characteristics).

The GMS does not report any overall judgment on the methodological quality of studies included in the MTCs. The ability of the GMS (or the ERG) to judge the similarity of trial populations is hampered by variable reporting of baseline characteristics that may be considered likely to be influential. While the GMS presents tabulations of mean age, baseline disability, previous relapse rate, proportion of patients having previous treatment and mean disease duration, these include substantial missing data as these variables have not been reported for all trials. In addition socio-demographic information and comorbidity were not reported further hampering the comparison of baseline characteristics between included trials. The GMS considers two main aspects of heterogeneity:

1. Pre-analysis considerations of scope of search/ appropriateness of inclusion of categories of trial or outcome. Section 6.7.1 of the GMS provides a rationale for excluding trials recruiting patients before 2000 (due to changes in diagnostic criteria) and also trials with less than 80% of patients having RRMS.

2. Post-analytical consideration of results from tests for heterogeneity in which the GMS concludes that there is low evidence of statistical heterogeneity. This statement appears to be justified on the basis of the reported statistics.

There is no discussion of the appropriateness of pooling the results for trials of treatment-naïve and treatment-experienced patients despite the acknowledgement in bullet points on page 176 of the GMS that there is limited information on previous treatment (reported in 15 of the 30 included trials) and that only three included trials (two of which included alemtuzumab) were for treatment-naïve patients. Advice from clinical experts suggests that both of these populations are likely to be the types of participants eligible for alemtuzumab in clinical practice.

The exclusion of trials recruiting patients before 2000, on the basis of changes in diagnostic criteria and the observed decline in ARRs in clinical trials over time[12] has an unfortunate effect by excluding all the direct evidence (versus placebo) for SC IFNβ1-a and IFNβ1-b. Since the economic model requires estimates of relative treatment effect, compared with placebo, the majority of comparisons (SC IFNβ1-a, IFNβ1-b and IM IFNβ-1a in addition to alemtuzumab) entering the model have needed to be constructed using the MTC and have not been informed by any direct evidence. Comparing the base case and all trials MTC analyses with the available direct evidence for three month SAD indicates some inconsistencies (Table 5). As would be expected the key differences relate to the hazard ratios estimated for SC IFNβ1-a and IFNβ1-b. For the former the credible range is reduced in the all trials analysis (compared with the base case) with the upper limit of the 95% credible interval no longer exceeding one. For the latter, a statistically non-significant harm in the base case is reversed to a small statistically non-significant benefit.

An alternative approach would have been to conduct the MTCs for all outcomes, including all the relevant evidence, while controlling for relapse rate.

**Table 5 Comparison of three month SAD hazard ratios from base case and all trial MTC versus direct evidence**

| | Base case MTC | All trials MTC | Direct |
|---|---|---|---|
| Alemtuzumab 12 mg | [redacted] | [redacted] | None |
| IFNβ-1b 250 µg | [redacted] | [redacted] | 0.68 (0.4 – 1.17) |
| Intramuscular (IM) IFNβ-1a | [redacted] | [redacted] | None |

Superseded – see Erratum

| | | | |
|---|---|---|---|
| SC IFNβ-1a | [redacted] | [redacted] | 0.65 (0.45 – 0.94) |
| GA 20 mg | [redacted] | [redacted] | 0.93 (0.65 – 1.32) |
| Fingolimod 0.5 mg | [redacted] | [redacted] | 0.76 (0.61 – 0.94) |
| Natalizumab 300 mg | [redacted] | [redacted] | 0.58 (0.28 – 1.19) |
| Note: in addition the trials of SC IFNβ1-a versus placebo and of IFNβ1-b versus placebo that were excluded for being recruiting patients prior to 2000, two trials reporting relevant data for GA 20 mg versus placebo were also excluded from the base case MTC | | | |

The ERG have checked the WinBUGS code submitted with the GMS in Additional Appendix 8 and note that it does not appear to take account of the inclusion of multi-arm trials (BEYOND, CAMMS223, CONFIRM, DEFINE, FREEDOMS, SELECT, TEMSO, TENERE, TOWER and TRANSFORMS all have more than two arms included in the analyses). Estimates of relative treatment effects from trials with more than two treatment arms will be correlated and analyses based on estimates of relative treatment effect (for example, HRs as used in the analysis of 3 month and 6 month SAD) should take this into effect. The ERG re-ran the analyses adopting a method for addressing this problem suggested by Woods and colleagues.[13] This had a limited impact on the results (see Table 6).

**Table 6 Three month SAD HRs from MTC – ERG analysis**

| | Alemtuzumab 12 mg | Placebo |
|---|---|---|
| Alemtuzumab 12 mg | NA | [redacted] |
| IFNβ-1b 250 µg | [redacted] | [redacted] |
| Intramuscular (IM) IFNβ-1a | [redacted] | [redacted] |
| SC IFNβ-1a | [redacted] | [redacted] |
| GA 20 mg | [redacted] | [redacted] |
| Fingolimod 0.5 mg | [redacted] | [redacted] |
| Natalizumab 300 mg | [redacted] | [redacted] |

The inclusion of data from all arms of multi-arm trials studies in the MTC is inconsistent, with all dosages of alemtuzumab from CAMMS223[1] included, but only one (alemtuzumab 12mg) from CARE-MS II.[3]

Overall the approach to the MTCs presented in the GMS appears to be reasonable and is the only viable method to provide comparisons between all comparators indicated in the scope developed by NICE and to populate the economic model adopted for the appraisal. However, it needs to be borne in mind that the analyses have required the combination of trials, many of which have reported limited baseline characteristics and which appear to differ substantially in

Superseded – see Erratum

the prior treatments received by patients. The analyses undertaken in the MTCs have not controlled for any of the differences in patient characteristics other than to exclude trials recruiting patients prior to 2000 and to test the robustness of this exclusion criterion (and the inclusion of trials recruiting a proportion of non-RRMS patients) through sensitivity analysis. It should also be noted that the economic model requires that the MTCs derive estimates of relative treatment effects for comparators against placebo, the majority of which (SC IFNβ1-a, IM IFNβ1-a, IFNβ1-b and Alemtuzumab) have not been studied in placebo-controlled trials (IM IFNβ1-a and Alemtuzumab) or where the placebo controlled studies were excluded from the base case MTC (SC IFNβ1-a and IFNβ1-b).

### 3.2 Summary statement of manufacturer's approach to evidence synthesis

The ERG assessed the quality of the GMS based on CRD questions for a systematic review and a summary of the overall quality of the submission can be seen in Table 7.

**Table 7 Quality assessment (CRD criteria) of MS review**

| **CRD Quality Item: score Yes/ No/ Uncertain with comments** | |
|---|---|
| 1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question? | 1. Yes |
| 2. Is there evidence of a substantial effort to search for all relevant research? i.e. all studies identified | 2. Yes |
| 3. Is the validity of included studies adequately assessed? | 3. Partly. Not all of the assessments were presented in the main report. The CRD quality assessment was completed for all 3 RCTs from the review (GMS page 109). The quality assessment for the MTC RCTs was in GMS Appendix 5 (page 418) and that of the 2 included non-RCTs in GMS Appendix 7 (page 432). One question did not appear to be assessed appropriately as it stated 'no comment'. No discussion of the effects of key biases was discussed in the text. |
| 4. Is sufficient detail of the individual studies presented? | 4. Partly. Not all the outcome data is reported comprehensively, with additional information presented in appendices. |
| 5. Are the primary studies summarised appropriately? | 5. Yes. |

The processes for inclusion/exclusion and data extraction were reported in the GMS and were assessed as being adequate by the ERG. There is less detail provided for the processes of undertaking the quality assessment. The majority of the information about the processes for the inclusion/exclusion and data extraction was found in the appendices rather than the main submission.

The ERG considers the submitted evidence to partly reflect the defined decision problem in the GMS. The main issue is the relation of the populations in the included trials to the scoped population of people with RRMS as described above, in particular with respect to the scoped subgroups.

Overall the chance of any systematic error in the systematic review based on the methods employed is uncertain.

### 3.3 Summary of submitted evidence

The ERG has reproduced data for the key outcomes from the included trials, the direct meta-analysis results and the MTC. For some of the outcomes (including the pooled data) the GMS report these as RR and the ERG have followed this convention. However, on checking the data inputs it would appear that for some outcomes (proportion relapse free and discontinuations) odds ratios were used, and for others (ARR and SAD) hazard ratios were used.

**Summary of results for SAD**

SAD at 3 months

Table 8 shows the results for SAD at three months. In the treatment-naïve participants in the CAMMS223 trial,[1] SAD at three months [redacted]. In the CARE-MS I trial,[2] (also treatment-naïve participants) there was [redacted]. There was also [redacted] on SAD at three months in the CARE-MS II[3] trial in those who had previously been treated with IFNβ-1a.

**Table 8: Sustained accumulation of disability at 3 months**

| **CAMMS223[1]** (treatment naïve participants) | | | | |
|---|---|---|---|---|
| | **IFNβ-1a (n=111)** | **Alemtuzumab 12mg (n=112)** | **HR vs IFNβ-1a (95% CI), p value** | **Treatment effect (95% CI)** |
| Patients with event, n (%) | 30 (32.7) | 16 (16.3) | [redacted] | [redacted] |
| **CARE-MS I[2]** (treatment naïve participants) | | | | |
| | **IFNβ-1a (n=187)** | **Alemtuzumab 12mg (n=376)** | **RR vs IFNβ-1a (CI), p value** | **Treatment effect (CI)** |
| Patients with event, n (%) | NR | NR | [redacted] | [redacted] |
| **CARE-MS II[3]** (previously treated participants) | | | | |
| | **IFNβ-1a (n=202)** | **Alemtuzumab 12mg (n=426)** | **RR vs IFNβ-1a (CI), p value** | **Treatment effect (CI)** |

| | | | | |
|---|---|---|---|---|
| Patients with event, n(%) | NR | NR | [redacted] | [redacted] |

NR, not reported. [a]Calculated by reviewer.

The pooled RR of 3-month SAD for alemtuzumab versus IFNβ-1a from the three RCTs was [redacted]

Results from the three alemtuzumab trials compared with the other relevant comparators via the MTC can be seen in Table 9. This shows that the comparisons of alemtuzumab with [redacted] were statistically significantly different. The comparisons between alemtuzumab [redacted], alemtuzumab and [redacted], and alemtuzumab and [redacted] were not statistically significant. None of the IFN-β treatments or GA were better than placebo.

**Table 9: Sustained accumulation of disability at 3 months from the MTC**

| | **Alemtuzumab 12 mg, RR (95% CrI)** | **Placebo, RR (95% CrI)** |
|---|---|---|
| Alemtuzumab 12 mg | [redacted] | [redacted] |
| IFNβ-1b 250 µg | [redacted] | 1.21 (0.68, 2.16)[a] |
| IM IFNβ-1a 30 µg | [redacted] | 0.91 (0.61, 1.33)[a] |
| SC IFNβ-1a 44 µg | [redacted] | 0.79 (0.51, 1.24)[a] |
| GA 20 mg | [redacted] | 0.93 (0.59, 1.45)[a] |
| Fingolimod 0.5 mg | [redacted] | 0.75 (0.58, 0.96) |
| Natalizumab 300 mg | [redacted] | 0.58 (0.4, 0.84) |

[a] Not statistically significant

<u>SAD at 6 months (co-primary endpoint)</u>

Table 10 shows the results for SAD at six months. In the treatment naïve participants in the CAMMS223 trial,[1] SAD at six months was statistically significantly better in those treated with 12 mg alemtuzumab than those treated with IFNβ-1a. In the CARE-MS I trial,[2] (also treatment naïve participants) there was no statistically significant treatment effect shown. In those previously treated with IFNβ-1a in the CARE-MS II[3] trial, a statistically significant treatment effect on SAD at six months was seen.

**Table 10: Sustained accumulation of disability at 6 months**

| **CAMMS223[1]** (treatment naïve participants) | | | | |
|---|---|---|---|---|
| | **IFNβ-1a (n=111)** | **Alemtuzumab 12mg (n=112)** | **HR vs IFNβ-1a (95% CI), p value** | **Treatment effect (95% CI)** |
| Patients with event, n (%) | 24 (26.2) | 8 (8.5) | 0.25 (0.11 to 0.57), p<0.001 | 75%(43 to 89) |

| **CARE-MS I[2]** (treatment naïve participants) | | | | |
|---|---|---|---|---|
| | **IFNβ-1a (n=187)** | **Alemtuzumab 12mg (n=376)** | **HR vs IFNβ-1a (CI), p value** | **Treatment effect (CI)** |
| Patients with event, n (%) | 20 (11%) | 30 (8%) | 0.70 (0.40 to 1.23), p=0.22 | 30% |
| **CARE-MS II[3]** (previously treated participants) | | | | |
| | **IFNβ-1a (n=202)** | **Alemtuzumab 12mg (n=426)** | **HR vs IFNβ-1a (CI), p value** | **Treatment effect (CI)** |
| Patients with event, n (%) | 40 (20%) | 54 (13%) | 0.58 (0.38 to 0.87), p=0.0084 | 42% |

Note analyses for 6-month SAD were based on per-protocol set for CAMMS223[1] and CARE-MS II.[3]

The pooled RR for SAD at 6 months for alemtuzumab versus IFNβ-1a from the three RCTs was [redacted]

Results from the three alemtuzumab trials compared with the other relevant comparators via the MTC can be seen in Table 11. This shows that the comparison between alemtuzumab and [redacted] was statistically significantly different. [redacted]

**Table 11: Sustained accumulation of disability at 6 months from the MTC**

| | **Alemtuzumab 12 mg, RR (95% CrI)** | **Placebo, RR (95% CrI)** |
|---|---|---|
| Alemtuzumab 12 mg | [redacted] | [redacted] |
| IFNβ-1b 250 µg | [redacted] | [redacted] |
| IM IFNβ-1a 30 µg | [redacted] | [redacted] |
| SC IFNβ-1a 44 µg | [redacted] | [redacted] |
| GA 20 mg | [redacted] | [redacted] |
| Fingolimod 0.5 mg | [redacted] | [redacted] |
| Natalizumab 300 mg | [redacted] | [redacted] |

[a] Not statistically significant

**Summary of results for relapse**

Relapse rate (co-primary endpoint)

In the three included RCTs,[1-3] treatment with alemtuzumab statistically significantly reduced the rate of relapse when compared with IFNβ-1a (Table 12). The HR in CAMMS223[1] was 0.31, the RRs in the two CARE-MS trials[2;3] ranged from 0.45 to 0.51. Follow-up was three years in the CAMMS223[1] trial and two years each in the CARE-MS I[2] and CARE-MS II[3] trial.

**Table 12: Relapse rate, annualised relapse rate, and proportion relapse free**

| **CAMMS223**[1] (treatment naïve participants) | | | | |
|---|---|---|---|---|
| **Relapse rate** | **Alemtuzumab 12mg (n=112)** | **IFNβ-1a (n=111)** | **HR vs IFNβ-1a (CI), p value** | **Treatment effect (CI)** |
| Total number of events<br>Patients with any event, n (%) | 34<br>24 | 89<br>45 | 0.31 (0.18 to 0.52), p <0.001 | 69% (48 to 82) |
| | **Alemtuzumab 12mg (n=112)** | **IFNβ-1a (n=111)** | **p value** | |
| ARR (95% CI) | 0.11 (0.08 to 0.16) | 0.36 (0.29 to 0.44) | NR | |
| Patients relapse free, %[a] | 77.0 | 51.6 | p<0.001 | |
| **CARE-MS I**[2] (treatment naïve participants) | | | | |
| | **Alemtuzumab 12mg (n=376)** | **IFNβ-1a (n=187)** | **RR vs IFNβ-1a (CI), p value** | **Risk reduction** |
| Total number of events<br>Patients with any event, n(%) | 119<br>82 (22) | 122<br>75 (40) | 0.45 (0.32 to 0.63), p<0.0001 | 54.9% |
| | **Alemtuzumab 12mg (n=376)** | **IFNβ-1a (n=187)** | **p value** | |
| ARR (95% CI) | 0.18 (0.13 to 0.23) | 0.39 (0.29 to 0.53) | NR | |
| Patients relapse free, % (95% CI)[a] | 77.6 (72.9 to 81.6) | 58.7 (51.1 to 65.5) | <0.0001 | |
| **CARE-MS II**[3] (previously treated participants) | | | | |
| | **Alemtuzumab 12mg (n=426)** | **IFNβ-1a (n=202)** | **RR vs IFNβ-1a (CI), p value** | **Risk reduction** |
| Total number of events<br>Patients with any event, n(%) | 236<br>147 (35%[a]) | 201<br>104 (53%[a]) | 0.51 (0.39 to 0.65), p<0.0001 | 49.4% |
| | **Alemtuzumab 12mg (n=426)** | **IFNβ- 1a (n=202)** | **p value** | |
| ARR (95% CI) | 0.26 (0.21 to 0.33) | 0.52 (0.41 to 0.66) | NR | |
| Patients relapse free, % (95% CI)[a] | 65.4% (60.7-69.7) | 46.7% (39.5-53.5) | p<0.0001 | |

[a] Kaplan Meier estimation. Note analyses for relapse rates were based on per-protocol set for CAMMS223[1] and CARE-MS II.[3]

ARR

The ARR from the three trials in the CAMMS223 trial[1] was 0.11 in the alemtuzumab 12 mg treated arm and 0.36 in the IFNβ-1a treated arm (see Table 12). In the CARE-MS I trial[2] the rates for alemtuzumab 12 mg and the IFNβ-1a groups respectively were 0.18 and 0.39. ARR in the CARE-MS II trial[3] was 0.26 in the alemtuzumab arm and 0.52 in the IFNβ-1a group (Table 12).

The pooled RR of ARR for alemtuzumab versus IFNβ-1a from the three RCTs was [redacted]

Results from the three alemtuzumab trials compared with the other relevant comparators via the MTC can be seen in Table 13. This shows that the comparison of alemtuzumab with [redacted] were statistically significantly different. The comparison between alemtuzumab and [redacted] was not statistically significant.

**Table 13: ARR from the MTC**

| | **Alemtuzumab 12 mg, RR (95% CrI)** | **Placebo, RR (95% CrI)** |
|---|---|---|
| Alemtuzumab 12 mg | [redacted] | [redacted] |
| IFNβ-1b 250 µg | [redacted] | 0.68 (0.52, 0.88) |
| IM IFNβ-1a 30 µg | [redacted] | 0.78 (0.67, 0.91) |
| SC IFNβ-1a 44 µg | [redacted] | 0.62 (0.51, 0.76) |
| GA 20 mg | [redacted] | 0.64 (0.53, 0.76) |
| Fingolimod 0.5 mg | [redacted] | 0.46 (0.4, 0.54) |
| Natalizumab 300 mg | [redacted] | 0.31 (0.25, 0.39) |

[a] Not statistically significant.

Proportion relapse free

Table 12 shows the proportion of participants classified as relapse free from the three alemtuzumab trials. In all three trials, alemtuzumab led to statistically significantly greater proportions of participants being relapse free than did IFNβ-1a. In the two trials[1;2] in treatment naïve participants, the proportions of relapse free were around 77% in the alemtuzumab groups and ranged from around 52% to 59% in the IFNβ-1a groups. In the trial[3] with previously treated participants, the proportion relapse free was 65.4% in the alemtuzumab group and 46.7% in the IFNβ-1a group.

The pooled RR of proportion relapse free for alemtuzumab versus IFNβ-1a from the three RCTs was [redacted]

Results from the three alemtuzumab trials compared with the other relevant comparators via the MTC can be seen in Table 14. This shows that

[redacted]

**Table 14: Proportion relapse-free from the MTC**

| | **Alemtuzumab 12 mg, RR (95% CrI)** | **Placebo, RR (95% CrI)** |
|---|---|---|
| Alemtuzumab 12 mg | [redacted] | [redacted] |
| IFNβ-1b 250 µg | [redacted] | [redacted] |
| IM IFNβ-1a 30 µg | [redacted] | [redacted] |
| SC IFNβ-1a 44 µg | [redacted] | [redacted] |
| GA 20 mg | [redacted] | [redacted] |
| Fingolimod 0.5 mg | [redacted] | [redacted] |
| Natalizumab 300 mg | [redacted] | [redacted] |

[a] Not statistically significant.

**Summary of results for EDSS**

The three included trials[1; 2; 3] all reported the change in mean EDSS score, and results can be seen in Table 15. Caution is required in the interpretation of these data owing to the limitations of the EDSS.[5] In the two trials with treatment naïve participants one demonstrated an improvement in EDSS score in the alemtuzumab treated group and a deterioration in the IFNβ-1a treated group (no p-value reported).[1] The other trial[2] found no differences between the two groups, with both groups showing a slight improvement. In the trial[3] in participants who had been previously treated, alemtuzumab 12 mg led to an improvement in mean EDSS compared with a deterioration with IFNβ-1a , and the difference was statistically significant.

**Table 15: change in mean EDSS score**

| **CAMMS223[1]** (treatment naïve participants) | | | | |
|---|---|---|---|---|
| **Mean (95% CI)** | **Alemtuzumab 12mg (n=112)** | **IFNβ-1a (n=111)** | **Difference between groups** | **p-value** |
| | -0.32 (−0.55 to −0.10) | 0.38 (0.13 to 0.63) | 0.7[a] | NR |
| **CARE-MS I[2]** (treatment naïve participants) | | | | |
| **Mean (95% CI)** | **Alemtuzumab 12mg (n=376)** | **IFNβ-1a (n=187)** | **RR vs IFNβ-1a (CI), p value** | **p-value** |
| | -0.14 (-0.25 to -0.02) | -0.14 (-0.29 to 0.01) | 0 | 0.97 |
| **CARE-MS II[3]** (previously treated participants) | | | | |
| **Mean (95% CI)** | **Alemtuzumab 12mg (n=426)** | **IFNβ-1a (n=202)** | **RR vs IFNβ-1a (CI), p value** | **p-value** |
| | -0.17 (-0.29 to -0.05) | 0.24 (0.07 to 0.41) | 0.41 | <0.001[b] |

EDSS scores range from 0 to 10, with higher scores indicating worse function. [a] Calculated by reviewer. [b] Also reported as p<0.0001.

**Summary of Health related quality of life (HRQoL)**

HRQoL was assessed using the SF-36 and the FAMS, although limited results were reported. The GMS also states that the EQ-5D was a tertiary outcome, but no data were reported.

CARE-MS I[2] and II[3] both reported the change from baseline for the FAMS for year 1 and 2 in the narrative of the GMS. This suggested that alemtuzumab patients had statistically significantly greater improvements from baseline than IFNβ-1a patients in both studies (p values only reported). The GMS refers to two published abstracts but do not represent their data on QoL. In CARE-MS I (Gupta 2012[14]) the difference in the mean change from baseline FAMS between treatment groups favoured alemtuzumab 12 mg/day (estimated by the ERG to be 3.85 at six months, 4.41 at 12 months, 5.16 at 18 months and 4.25 at 24 months). In CARE-MS II (Arroyo 2013[15]) the estimated mean differences in FAMS between treatment groups were reported to be 5.0 points (95% CI: 1.8, 8.2) at six months, 5.6 (95% CI 2.2, 9.1) at 12 months, 7.8 (95% CI 4.0, 11.6) at 18 months and 5.3 (95% CI 1.3, 9.4) at 24 months. Clinical advice to the ERG suggests that these differences are around the region they would suggest to be clinically meaningful.

All three RCTs also reported the change from baseline for the SF-36, [redacted]. CARE-MS I[2] and II[3] reported the SF-36 outcome for year 1 and CARE-MS II for year 2, and the GMS states that the only the physical component summary score was statistically significantly improved with alemtuzumab. The GMS refer to an abstract (Selmaj 2012[16]) which the ERG have checked and no further data were available for the SF-36. The abstract states that on the EQ-VAS, alemtuzumab patients (in both CARE-MS trials[2;3]) improved significantly more than IFNβ patients at month six and 12 (p-values <0.05), but not at 18 or 24 months.

In summary: In the co-primary outcome of 6 month SAD a statistically significant effect was seen in two of the three trials. In the pooled comparison and MTC results [redacted]. The CAMMS223[1] trial is likely to be contributing the biggest treatment effect to these comparisons. The co-primary outcome of relapse rate appears to be more consistent across the three trials.

**Sub-group analyses results**

Both the CARE MS I[2] and II[3] trials conducted subgroup analyses to assess the influence of baseline or demographic factors on relapse rate. In addition, alemtuzumab's effect on subgroups defined by previous therapy or anti-interferon antibodies (either present at baseline or emerging subsequently) was reported by CARE MS II.[3] However, it is unclear if these analyses were defined *a priori* and these have not been summarised here.

In the CAMMS223[1] and CARE-MS II[3] trials subgroup analyses were undertaken of participants with RES RRMS (GMS page 125). The analysis was of RES RRMS and this was defined as ≥2 relapses in the year prior to treatment and at least 1 gadolinium-enhancing lesion at baseline. This differs slightly to the definition of RES RRMS in the licensed indication and NICE guidance for natalizumab (NICE 2007) *(≥2 disabling relapses in one year with ≥1 gadolinium-enhancing lesions on brain MRI or a significant increase in T2 lesion load as compared to a previous recent MRI*.) The subgroup analysis was defined *a priori* for some outcomes but was post-hoc for others (see GMS page 102) and ERG report p27.

No data were reported on the number of participants falling within this subgroup in the GMS and results were presented in a summary table only (B6.5.6, page 125). Key results for the two outcome analyses that were pre-planned (relapse and SAD) are reproduced below. Caution is required in the interpretation of these data, as in most cases the alemtuzumab groups were pooled. The GMS refers to three published abstracts which the ERG has also checked. These abstracts provide some information from the CAMS 223[1] trial and CARE-MS II[3] trial. The proportions of participants reported to meet the criteria for having highly active disease at baseline in CAMS223 (Wingerchuk (2010)[17] were reported as 125 (56.3%) of alemtuzumab participants and 61 (55%) of IFNβ-1a participants. Krieger (2013)[18] and Confayreux (2012)[19] state that 101 (23.7%) of alemtuzumab and 42 (20.8%) of IFNβ-1a participants had highly active RRMS at baseline in CARE-MS II.[3]

<u>Relapse and ARR</u>

In the CAMMS223[1] trial of participants who were treatment naïve the results for relapse for this subgroup were from the two alemtuzumab groups (12 mg and 24 mg) combined. The GMS suggests that there was an 81% lower rate of relapse with alemtuzumab compared with IFNβ-1a ($p<0.0001$). No other data were reported. ARR was reported to be 0.09 in the alemtuzumab 12 mg subgroup and the GMS stated this was statistically significantly lower than IFNβ-1a ($p<0.005$). No data were reported for the IFNβ-1a group ARR, however in the Wingerchuk[17] abstract this was reported to be 0.47.

In the CARE-MS II[3] trial (in a previously treated population) the GMS reports a 51% reduction in ARR to year two in the alemtuzumab 12 mg subgroup compared with the IFNβ-1a subgroup (0.33 versus 0.65 respectively, no p-value reported). There was also reported to be a 56%

reduction in the proportion of participants with relapse in the alemtuzumab 12 mg subgroup compared with the IFNβ-1a subgroup (p=0.0018, no other data reported).

In the two published abstracts from CARE-MS II[3] relapse after 2 years was reported to be 35.8% for the subgroup in the alemtuzumab arm, and 60% in the IFNβ-1a arm (Krieger[18]). Over two years the ARR was reported to be 0.33 (95% CI: 0.24-0.46) for alemtuzumab subgroup and 0.65 (0.47-0.90) for IFNβ-1a subgroup (Confayreux[19]). No p-values were reported in these abstracts.

SAD

The GMS reported that in the CAMMS223[1] trial, the estimated proportion of participants in the two alemtuzumab subgroups with SAD at six months was 91% (95% CI 84.0, 95.2) compared with 73% (95% CI 58.1, 83.5) in the IFNβ-1a subgroup, HR 0.30 (95% CI 0.13, 0.69), p=0.0045. The GMS stated that comparisons of each dose group to IFNβ-1a were similar and that with the 12 mg alemtuzumab subgroup, there was a 65% lower risk of SAD at six months (p=0.36).

In the CARE-MS II[3] trial, there was a 51% reduction difference in the percentage of patients with SAD Years 0 - 2 in the alemtuzumab RES patient subgroup (8.95%) compared to SC IFNβ-1a RES patients (17.62%). The GMS reported a 77% increase in the number of patients with sustained reduction of disability in the alemtuzumab treatment group compared to the SC IFNβ-1a group (22.99% vs. 12.99%).

No subgroup analyses were undertaken on those with highly active RRMS who had been previously treated with DMTs. No subgroup analyses were formally presented for the two other subgroups defined in the NICE scope (treatment naive and previously treated populations), however, two trials reported in the GMS reflect the former group and one the latter group and results shown for these trials could be used to consider these populations. Results are seen in Section 3.3.

**Summary of adverse events**
**Adverse Events (AEs) – Alemtuzumab**

Safety data and adverse events in the GMS were presented through to 31 December 2011, with additional data from the Safety Update Report (SUR) which provided updated data from the ongoing extension study through to 26 November 2012 (CAMMS03409[6]). The majority of the

Superseded – see Erratum

data were CIC. Data for AEs were based on the same three trials as included in the systematic review and meta-analysis, but were pooled with the addition of the CAMMS03409[6] extension study. A summary of the safety overview for the three RCTs was presented in Additional Appendix 9.

Few AE data for IFNβ-1a were reported in the main GMS. Data for Grade 3 – 5 were extracted by the ERG from the Additional Appendix 9 and are discussed in the section below.

AEs were reported in all patients (100% in both treatment arms) at year 3 in the CAMMS223[1] trial, in 92.0% of IFNβ-1a and 96.0% alemtuzumab-treated of patients at year 2 in the CARE-MS I[2] trial, and in 94.6% IFNβ-1a and 98.4% alemtuzumab-treated of patients at year 2 in the CARE-MS II[3] trial (GMS page 27). Table 16 provides an overview of the pooled AEs data in alemtuzumab-treated patients.

**Table 16: Overview of AEs in alemtuzumab-treated patients**

| AEs | Alemtuzumab 12 mg, (n=1217)[a] |
|---|---|
| Any Event | [redacted] |
| Related | [redacted] |
| Unrelated | [redacted] |
| Grade 1 | [redacted] |
| Grade 2 | [redacted] |
| Grade 3 | [redacted] |
| Grade 4 | [redacted] |
| Grade 5 | [redacted] |
| AEs leading to treatment withdrawal | [redacted] |
| AEs leading to study discontinuation | [redacted] |

[a] Cumulative up to and including 26/11/2012. AE intensity was graded as - Grade 1: mild; Grade 2: moderate; Grade 3: severe; Grade 4: very severe; Grade 5: fatal.

The incidence of AEs for all alemtuzumab 12 mg-treated patients (pooled data) over all available follow-up was [redacted](cumulative up to and including 31/12/2011). Overall incidence of AEs for patients treated with alemtuzumab declined [redacted]. Subgroup data per the number of treatment courses given were not presented, which may have been more representative of incidence rates and explain the decline. The GMS (page 191) cites skin and subcutaneous tissue disorders [redacted]), nervous system disorders [redacted] and infections and infestations [redacted] as the three most frequently affected MedDRA system organ

classes, and [redacted], pyrexia (30.9%), and MS relapse [redacted] as the most common AEs (pooled data). The incidence of thyroid AEs increased from [redacted] in year 1 to [redacted] (GMS page 192). The GMS suggests that this is due to higher incidences of [redacted] Table 17 presents grouped AEs reported at a ≥5% incidence in alemtuzumab-treated patients (pooled data) as this was relevant to the economic evaluation.

**Table 17: Grouped AEs reported at a ≥5% incidence in alemtuzumab-treated patients**

| **Grouped AEs** | **Alemtuzumab 12 mg, (n=1217)[a]** |
|---|---|
| Blood and lymphatic system disorders | [redacted] |
| Cardiac disorders | [redacted] |
| Ear and labyrinth disorders | [redacted] |
| Endocrine disorders | [redacted] |
| Eye disorders | [redacted] |
| Gastrointestinal disorders | [redacted] |
| General disorders and administration site conditions | [redacted] |

[a] Cumulative up to and including 26/11/2012. AE intensity was graded as - Grade 1: mild; Grade 2: moderate; Grade 3: severe; Grade 4: very severe; Grade 5: fatal.

Adverse events (pooled data) led to 2.6% of alemtuzumab-treated patients withdrawing from treatment and 0.3% discontinuing with the trial (see Table 16). [redacted]

**Table 18: Discontinuations due to adverse events odds ratio results from the MTC**

| | Alemtuzumab 12 mg, Rate ratio (95% CrI) | Placebo, Rate ratio (95% CrI) |
|---|---|---|
| Alemtuzumab 12 mg | [redacted] | [redacted] |
| IFNβ-1b 250 µg | [redacted] | [redacted] |
| IM IFNβ-1a 30 µg | [redacted] | [redacted] |
| SC IFNβ-1a 44 µg | [redacted] | [redacted] |
| GA 20 mg | [redacted] | [redacted] |
| Fingolimod 0.5 mg | [redacted] | [redacted] |
| Natalizumab 300 mg | [redacted] | [redacted] |

* Not statistically significant

**Serious Adverse Events (SAEs) – Alemtuzumab**

The incidence of SAEs through to 31 December 2011 from the pooled studies was [redacted] with [redacted] the most frequently reported

Superseded – See Erratum

MedDRA system organ classes for the alemtuzumab 12 mg group. [redacted] The GMS provided a full list of all SAEs in ≥2 alemtuzumab-treated patients (any group) over all available follow-up in Additional Appendix 9.

In the pooled studies incidence of SAEs decreased over time ([redacted] but the rate of SAEs by number of treatment courses received was [redacted] in the alemtuzumab-treated group ([redacted]) and by [redacted]. This suggests [redacted] The GMS suggests that there appeared to be [redacted] over time (GMS page 192). With alemtuzumab related risks included infusion associated reactions (IARs), autoimmune disorders (thyroid, ITP, and nephropathies including anti-glomerular basement membrane) and infections (see Table 22).

In the pooled data:
*Anaphylactic reactions:* (defined according to the Sampson Standardised MedDRA Queries (SMQ) criteria) were identified in 7.7% of patients in the alemtuzumab 12 mg group (treatment cycle 1: 6.7%, cycle 2: 1.5%, cycle 3: 2.1%, cycle 4 and 5: 0%).

*Autoimmune disease*: (consisting of hyperthyroidism, hypothyroidism, Grave's disease and ITP) the incidence of autoimmune adverse events plateaued to a rate of approximately [redacted] in the trials after 5 years, starting from 2 weeks after initial treatment and most frequent 12–18 months after first treatment, with no new cases ≥60 months or more after initial treatment.

*Cytopenias:* (such as autoimmune hemolytic anaemia) occurred [redacted] of alemtuzumab-treated patients, with >5% of patients experiencing [redacted]. Serious cytopenia AEs were reported in [redacted] of patients, with incidence highest in years [redacted] and no serious cytopenias reported after year 4.

*Idiopathic thrombocytopenia purpura (ITP):* the GMS stated that this was the second most frequent autoimmune AE in alemtuzumab-treated patients, occurring in [redacted]) alemtuzumab-

Superseded – see Erratum

treated patients (12mg dose), with serious ITP AEs in [redacted] of patients. The most common onset occurred after [redacted] in years 1 to 3.

*Nephropathies*: alemtuzumab has been associated with glomerular disease, particularly Goodpasture's disease (anti-GBM disease) which can result in permanent lung and kidney damage and often death. Additional monitoring for anti-GBM disease was put in place following the identification of a case in the CAMMS223 trial. The GMS reports that nephropathies occurred in [redacted] of alemtuzumab-treated patients, with events occurring within 39 months following the last administration of alemtuzumab.

*Thyroid diseases*: the GMS stated that this was the most frequent autoimmune AE in alemtuzumab-treated patients, occurring in around 36.2% of patients during the 4 years after the first treatment course, with an increased risk between [redacted] months and the highest incidence in [redacted] after the first alemtuzumab treatment course. Serious thyroid events occurred in [redacted] of all alemtuzumab-treated patients over all available follow-up, with [redacted] requiring surgical treatment.

The manufacturer were asked for evidence of adverse event data by cycle and provided data for thyroid disorders, potential anaphylactic reactions, infusion associated reactions, and ITP (see clarification A10, page 11-12). The ERG have summarised key information in Table 19 - Table 21.

**Table 19: Thyroid Disorders in Alemtuzumab 12mg -Treated Patients (All Available Follow Up, Pool C)**

| | **Overall, n (%)** | **Cycle 1, n (%)** | **Cycle 2, n (%)** | **Cycle 3, n (%)** | **Cycle 4, n (%)** |
|---|---|---|---|---|---|
| Patients at Risk | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Any Thyroid AE | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Data presented are for all grades

**Table 20: Incidence of Infusion-Associated Reactions by Cycle and Severity in All Alemtuzumab 12mg -Treated Patients (All Available Follow Up, Pool C)**

| **System Organ Class** | **Overall, n** | **Cycle 1, n** | **Cycle 2, n** | **Cycle 3, n** | **Cycle 4, n** |
|---|---|---|---|---|---|
| Patients at risk | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

| | | | | | |
|---|---|---|---|---|---|
| Any IAR | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Grade 1 | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Grade 2 | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Grade 3 | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Grade 4 | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

[redacted]

**Table 21: First Immune Thrombocytopenic Purpura Event by Cycle in All alemtuzumab 12mg -Treated Patients, All Available Follow Up (Pool C)**

| Cycle | Alemtuzumab 12 mg/day (n=1216), n (%) |
|---|---|
| 1 | [redacted] |
| 2 | [redacted] |
| 3 | [redacted] |

Percentages are based on the number of treated patients meeting the platelet-based or AE-based definition of ITP in the corresponding treatment group. Data presented are for all grades

In addition, the GMS reported details reported for infections (GMS pages 196 -197) and malignancies (GMS pages 197-198).

**Table 22: SAEs rates of IARs, infections, thyroid and ITP per trial**

| | CAMMS223[1] (5 year follow-up) | CARE-MS I[2] (2 year follow-up) | CARE-MS II[3] (2 year follow-up) |
|---|---|---|---|
| IARs | [redacted] | [redacted] | [redacted] |
| Infections | [redacted] | [redacted] | [redacted] |
| Thyroid | [redacted] | [redacted] | [redacted] |
| ITP | [redacted] | [redacted] | [redacted] |

Based on all available follow-up in the pooled data, [redacted] of patients treated with alemtuzumab discontinued treatment due to an SAE, with [redacted] of patients discontinuing treatment due to IARs.

**Mortality**

[redacted]

**Adverse Events (AEs) – INFβ-1a**

AEs data for those treated with INFβ-1a were mostly reported in Additional Appendix 9. No pooled data combining the three trials were presented. Grade 3 AEs between the trials ranged from [redacted] and Grade 4 from [redacted], with [redacted] reported Grade 5 AEs (see Table 23). Withdrawals due to AEs ranged from 5.9% - [redacted] and discontinuations from [redacted] (not reported for CAMMS223[1]). The GMS suggests that rates were comparable between alemtuzumab and IFNβ-1a. However, in two out of the three RCTs, reported AEs were higher for alemtuzumab-treated patients.

**Table 23: Overview of AEs in SC IFNβ-1a-treated patients**

| **AEs - SC IFNβ-1a** | **CAMMS223[1]**<br>**n=107** | **CARE-MS I[2]**<br>**n=187** | **CARE-MS II[3]**<br>**n=202** |
|---|---|---|---|
| Grade 3 | [redacted] | [redacted] | [redacted] |
| Grade 4 | [redacted] | [redacted] | [redacted] |
| Grade 5 | [redacted] | [redacted] | [redacted] |
| AEs leading to treatment withdrawal | [redacted] | 5.9% | [redacted] |
| AEs leading to study discontinuation | Not reported | [redacted] | [redacted] |

AE intensity was graded as follows: grade 1: mild; grade 2: moderate; grade 3: severe; grade 4: very severe; grade 5: fatal, with only 3 – 5 are presented in the Table.

**Severe Adverse Events (SAEs) – INFβ-1a**

SAEs data for those treated with INFβ-1a were also only reported in Additional Appendix 9. No pooled data combining the three trials were presented. Grade 3 AEs between the trials ranged from [redacted] and Grade 4 from [redacted], with [redacted] reported Grade 5 AEs (see Table 24). Withdrawals due to AEs ranged from [redacted] and discontinuations from [redacted] (not reported for CAMMS223[1]). The GMS suggested that SAE rates were comparable between alemtuzumab and IFNβ-1a, which were higher for alemtuzumab-treated patients in one of the RCTs. SAEs were reported in 23.4% of the IFNβ-1a and 22.2% of the 12mg alemtuzumab-treated patients at year 3 in the CAMMS223[1] trial, 14.4% of the IFNβ-1a and 18.4% of the alemtuzumab-treated patients at year 2 in CARE-MS I[2] trial and [redacted] of the IFNβ-1a- and [redacted] of the alemtuzumab-treated patients at year 2 in CARE-MS II[3] trial.

Superseded - see Erratum

**Table 24: Overview of SAEs in SC IFNβ-1a-treated patients**

| SAEs - SC IFNβ-1a | CAMMS223[1] n=107 | CARE-MS I[2] n=187 | CARE-MS II[3] n=202 |
|---|---|---|---|
| Grade 3 | [redacted] | [redacted] | [redacted] |
| Grade 4 | [redacted] | [redacted] | [redacted] |
| Grade 5 | [redacted] | [redacted] | [redacted] |
| AEs leading to treatment withdrawal | [redacted] | [redacted] | [redacted] |
| AEs leading to study discontinuation | Not reported | [redacted] | [redacted] |

AE intensity was graded as follows: grade 1: mild; grade 2: moderate; grade 3: severe; grade 4: very severe; grade 5: fatal, with only 3 – 5 are presented in the Table.

# 4 ECONOMIC EVALUATION

## 4.1 Overview of manufacturer's economic evaluation

The manufacturer's submission to NICE includes:

i) a review of published economic evaluations of MS treatments compared for adults with RRMS or progressive MS (including Secondary Progressive MS [SPMS] or Primary Progressive MS [PRMS])

ii) a report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of alemtuzumab is compared with beta-interferons, glatiramer acetate, fingolimod and natalizumab for active RRMS.

**Manufacturer's review of published economic evaluations**

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of adults with RRMS or progressive MS (including SPMS or PRMS). See Section 3.1.1 of this report for the ERG critique of the search strategy. The review identified 33 studies evaluating cost effectiveness in MS, although none of these studies were of alemtuzumab.

**CEA Methods**

The cost effectiveness analysis (CEA) uses a multi-state Markov model to estimate the cost-effectiveness of alemtuzumab compared with other possible alternatives in adults with active RRMS. The model adopted a lifetime horizon of 50 years with a one year cycle length.

The model developed has health states based on disease classification (RRMS or SPMS) and severity (defined by the EDSS). The model was based on a structure developed by the School

of Health and Related Research (ScHARR) in the evaluation of beta-interferons for the treatment of MS.[20] Active RRMS patients entered the model with baseline characteristics collected for RRMS patients in the UK RSS.

Clinical data used in the model are based on results from the alemtuzumab trials and MTC for RRMS (in the base case), HA despite interferon use (in the subgroup analysis) and RES (in the subgroup analysis).

Patients transition through the model accounting for withdrawal, mortality, disease progression in terms of EDSS, conversion from RRMS to SPMS, and a DMT stopping rule as recommended by Association of British Neurologists guidelines.[21] Treatment effects are included in terms of 3-month SAD and ARR from the MTC. SAD HRs are applied to natural history transition matrices derived from the London Ontario active RRMS dataset and supplemented by the placebo arms of TOWER and TEMSO Treatment transition matrices are used to estimate progression of patients through the disease scale (EDSS) as well as the disease classification in terms of RRMS and SPMS.

Quality of life data used in the model accounted for EDSS level, whether a relapse had occurred, treatment-related adverse events and carer disutility.

Costs categories were based on the NHS and PSS perspective, and included treatment acquisition costs, administration costs, monitoring costs, adverse event costs, disease costs (associated with EDSS level), and relapse costs split by severity.

The model outputs in the GMS have been compared to published literature to validate outputs (GMS Section 7.7.1). The model results are compared for alemtuzumab versus IFN-β1a (Rebif) 44µg between clinical and model outputs at the end of year 2. There was reasonable agreement, with generally more relapses in the model than the trials, lower quality of life in the model than the clinical trial, and similar mortality.

**CEA Results**

Costs and QALYs per patient were discounted at 3.5% per annum. Results from the economic model are presented (section 7.7.6, page 337 of the GMS) as incremental cost per QALY

gained for alemtuzumab compared with beta interferons, glatiramer acetate, fingolimod and natalizumab.

For the base case an incremental cost per QALY gained of £8,924 versus glatiramer acetate is reported (see Table 25). The MS analysis shows that other relevant comparators were strongly or extendedly dominated. The deterministic sensitivity analysis indicated that the model was most sensitive to the 3-month SAD HR, when using the 95% upper and lower confidence intervals from the MTC. The model was also fairly sensitive to the inclusion criteria applied in the derivation of MTC results in terms of SAD, ARR and withdrawal. Scenario analysis showed that the model was also sensitive to assumptions around the waning of treatment effect.

**Table 25 Base case deterministic cost effectiveness results (GMS B7.7.12 page 338)**

| Technologies | Total costs (£) | Total QALYs | Inc costs (£)[a] | Inc QALYs [a] | ICER (£) versus baseline (QALYs)[a] | ICER (£) inc (QALYs) [b] |
|---|---|---|---|---|---|---|
| Glatiramer acetate | 487,842 | 2.745 | | | | |
| SC IFNβ-1a 44µg | 489,354 | 2.850 | 1,512 | 0.106 | 14,277 | Extendedly |
| SC IFNβ-1a 22µg | 490,388 | 2.854 | 2,545 | 0.110 | 23,227 | Extendedly dominated |
| IM IFNβ-1a | 494,626 | 2.764 | 6,784 | 0.019 | 354,272 | Dominated |
| Alemtuzumab | 499,347 | 4.034 | 11,505 | 1.289 | 8,924 | 8,924 |
| IFNβ-1b | 502,969 | 2.329 | 15,127 | -0.416 | Dominated | Dominated |
| Fingolimod (assumed PAS price £13,000) | 507,049 | 3.068 | 19,207 | 0.323 | 59,443 | Dominated |
| Fingolimod | 529,094 | 3.068 | 41,252 | 0.323 | 127,672 | Dominated |
| Natalizumab | 530,800 | 3.373 | 42,958 | 0.628 | 68,383 | Dominated |

ICER, incremental-cost effectiveness ratio. PAS, Patient access scheme. QALY, quality-adjusted life year.
[a] Results compared to glatiramer acetate, [b] Incremental analysis, results compared to next best option (that is not dominated or extendedly dominated)
Dominated: treatment is more costly and less effective than alternative treatment. Extendedly dominated: treatment produces additional gains in effectiveness at incremental costs higher than those of the next most effective strategy.

Probabilistic results are shown in Table B7.7.9 (GMS page 337). A cost-effectiveness acceptability curve (CEAC) containing all comparative treatments is shown in Figure B7.7.7, page 350. The probabilistic results indicate that alemtuzumab is cost-effective in the active RRMS population against all comparators above a WTP of approximately £7,100 per QALY gained.

The GMS also show analyses comparing alemtuzumab for HA RRMS versus fingolimod and RES RRMS versus natalizumab (Section 7.9, page 354). For both analyses, the comparative treatment were dominated by alemtuzumab, and the GMS concluded that alemtuzumab was the most cost-effective of all treatments in both HA and RES RRMS.

## 4.2 Critical appraisal of the manufacturer's submitted economic evaluation

**Manufacturer's review of published economic evaluations**

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of adults with RRMS or progressive MS (including SPMS or PRMS). The inclusion and exclusion criteria for the systematic review are listed in section 10.10 of the MS, page 439. The inclusion criteria state that economic evaluations of any intervention in MS would be included. The exclusion criteria state that patients with PPMS would be excluded.

58 studies were identified from screening 311 titles and abstracts. 33 studies were identified as evaluating cost effectiveness in MS. Of these cost-effectiveness studies, no studies for alemtuzumab were identified. However, two cost-effectiveness studies from the UK NHS perspective have been published for MS DMTs and used in NICE assessments.[20;22] The reasons for excluding full-text articles were: 23 Population: 1 Review: 1 News article: 1 Article Unobtainable: 20. The manufacturer has analysed and quality assessed the two cost-effectiveness studies using the Drummond checklist.[23] The manufacturer discussed the differences in methods between the two studies but does not discuss the study results.

The GMS does not discuss any systematic reviews of cost effectiveness studies for MS DMTs. The ERG conducted an ad hoc search and found three recent systematic reviews of cost effectiveness studies for MS DMT.[24-26] No review contained any cost effectiveness studies for alemtuzumab. The ERG considers it unlikely that any cost-effectiveness studies of rivaroxaban were missed by the manufacturer as the literature search methods appear sound. The ERG consequently did not re-run the cost-effectiveness search.

**Critical appraisal of manufacturer's submitted economic evaluation**

The ERG have considered the methods applied in the economic evaluation in the context of the critical appraisal questions listed in Table 26 below, drawn from common checklists for economic evaluation methods (e.g. Drummond and colleagues[23]). The critical appraisal

Superseded – see Erratum

checklist indicates that overall the manufacturer follows recommended methodological guidelines.

**Table 26: Critical appraisal checklist of economic evaluation**

| Item | Critical Appraisal | Reviewer Comment |
|---|---|---|
| Is there a well-defined question? | Y | As per NICE scope |
| Is there a clear description of alternatives? | Y | Alemtuzumab is compared to beta-interferons, glatiramer acetate, fingolimod and natalizumab as an alternative within active RRMS |
| Has the correct patient group / population of interest been clearly stated? | Y | HA despite treatment with beta interferon and RES subgroups are considered but not naïve and treatment experienced patients |
| Is the correct comparator used? | Y | |
| Is the study type reasonable? | Y | Cost utility analysis |
| Is the perspective of the analysis clearly stated? | Y | NHS and PSS |
| Is the perspective employed appropriate? | Y | |
| Is effectiveness of the intervention established? | Y | |
| Has a lifetime horizon been used for analysis (has a shorter horizon been justified)? | Y | |
| Are the costs and consequences consistent with the perspective employed? | Y | |
| Is differential timing considered? | Y | Discounted at 3.5% per annum for costs and benefits |
| Is incremental analysis performed? | Y | GMS Section 7.7.6 |
| Is sensitivity analysis undertaken and presented clearly? | Y | GMS Section 7.7.7 |

Y=Yes.

**NICE reference case**

The NICE reference case requirements have also been considered for critical appraisal of the submitted economic evaluation in Table 27. The submitted evaluation conforms with the NICE reference case.

**Table 27 NICE reference case requirements**

| NICE reference case requirements: | Included in submission | Comment |
|---|---|---|
| Decision problem: As per the scope developed by NICE | Y | |
| Comparator: Alternative therapies routinely used in the UK NHS | Y | Alemtuzumab is compared to beta-interferons, glatiramer |

| | | |
|---|---|---|
| | | acetate, fingolimod and natalizumab as an alternative within active RRMS |
| Perspective on costs: NHS and PSS | Y | |
| Perspective on outcomes: All health effects on individuals | Y | (Discussed in section 4.2) |
| Type of economic evaluation: Cost effectiveness analysis | Y | |
| Synthesis of evidence on outcomes: Based on a systematic review | Y | Outcomes derived from MTC. (Discussed in clinical evidence sections 3.3 and 3.4.) |
| Measure of health benefits: QALYs | Y | (Discussed in section 4.2) |
| Description of health states for QALY calculations: Use of a standardised and validated generic instrument | Y | (Discussed in section 4.2) |
| Method of preference elicitation for health state values: Choice based method (e.g. TTO, SG, not rating scale) | Y | (Discussed in section 4.2) |
| Source of preference data: Representative sample of the public | Y | (Discussed in section 4.2) |
| Discount rate: 3.5% p.a. for costs and health effects | Y | |

? = uncertain. N/A=not applicable. Y= Yes. P.A., per annum. SG, standard gamble, TTO, Time trade off

**Modelling approach / Model Structure**

The GMS economic model consisted of a multi-state Markov model with health states for EDSS, SPMS and death (Figure 1). Costs and QALYs were calculated over a life time horizon (50 years) and discounted at 3.5% per annum. Cost categories were based on the NHS and PSS perspective.

Patients enter the model in one of ten EDSS health states. In each annual cycle, active RRMS patients may remain in the same EDSS state, progress to a more severe EDSS state, convert to SPMS or die. Once a patient converts to SPMS, they may remain in the same EDSS state, progress to a more severe EDSS state or die. Death is represented by EDSS 10. The model also estimates the frequency of relapses leading to hospitalisation and not leading to hospitalisation.

The GMS uses two approaches to measure disease progression in the model. The first approach applied HRs, derived from a MTC, to a natural history dataset. This approach is used for the comparative analysis for all treatments and is referred to in the GMS as the “natural history comparison”.

The second approach used transition probabilities derived directly from patient level data of the two alemtuzumab trials (CARE-MS I[2] and CARE-MS II[3]). This method was used for a sensitivity analysis comparison between alemtuzumab and IFNβ-1a 44µg. This method is referred to as "direct comparison".

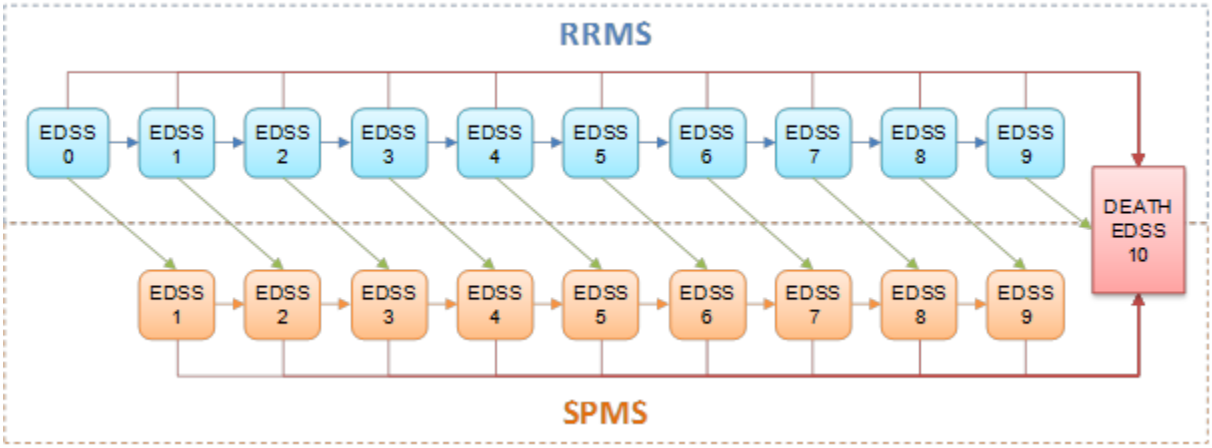


**Figure 1: Schematic of the multi-state Markov model**

The GMS states that the model structure allows estimation of the full impact of the disease across a lifetime, in terms of both costs and quality of life, from pre-diagnosis at EDSS 0 (normal neurological examination), to EDSS 9 (confined to bed) before reaching EDSS 10 (death). The GMS states that the model structure differs from the original ScHARR model[20] by capturing mortality separately from the transitions through EDSS states to allow for an increasing risk of mortality by age and gender.

The ERG considers that the structure of the model is consistent with currently accepted theory of MS. The structure of the model is based on the previously published ScHARR model, developed for NICE to evaluate the cost-effectiveness of beta-interferons and glatiramer acetate in the treatment of MS.[20] Adapted versions of the ScHARR model[20] have also been used in previous MS technology appraisals (TA).[27;28]

For the natural history comparison, the natural history of the disease was modelled based on real-world longitudinal observational disability progression data obtained from the London Ontario data set[29] (and placebo arms of TOWER and TEMSO for the transition probabilities from EDSS 0).[30] The original ScHARR model[20] (and subsequent technology appraisals) used the London Ontario data set.[29] The NICE TA for fingolimod raised concerns over this data set

and that it may have given more rapid disability progression rates than those seen in the clinical trials and in the current UK patient population.[27] The GMS conducted a systematic review to identify the most appropriate natural history transition matrix for disability progression for patients receiving no DMT (see GMS Additional Appendix 10). They were unable to find any more appropriate data and concluded that the London Ontario dataset was the most appropriate, robust and clinically plausible. The ERG considers that the manufacturer has not fully explored the uncertainty around the natural history of MS. In light of previous technology appraisals, the ERG suggests that alternative data sources should have been explored more extensively and validated, where possible, against trial data.

The progression of patients receiving DMTs is estimated by adjusting the natural history transition matrix by the relative effect of treatment versus placebo derived from the MTC. In addition, a relative risk was also applied for each treatment for the ARR and the risk of hospitalisation for those with relapses. A waning of treatment effect can be implemented for alemtuzumab patients according to the duration since starting treatment and for the treatment for those who transition to beyond EDSS ≥6 or SPMS.

Disease progression (as 3-month SAD) and relapses are modelled independently with independent treatment effects being applied to each (GMS page 278). The GMS justify this approach by stating that some treatments reduce relapses, others slow progression and modelling outcomes separately shows impact that the different costs and QALYs associated with a reduction in SAD and number of relapses have on cost-effectiveness of treatment.

All-cause mortality rates in the model were obtained from interim life tables for England and Wales from 2008-2010.[31] A weighted average all-cause mortality rate was calculated based upon the female to male ratio of MS patients (2.98:1).[32] These mortality rates were adjusted using mortality multipliers by MS disease severity from Pokorski and colleagues,[33] as previously used by TA 127[28] and TA 254.[27] The same mortality multipliers were applied to both RRMS and SPMS populations.

**Patient Group**

The patient group included in the economic evaluation is adults with active RRMS, defined by clinical or imaging features, which is in line with the expected marketing authorisation. The demographic profile at baseline is described and was sourced from the UK RSS[32] and is

therefore largely representative of UK RRMS population (although this scheme included patients with SPMS and therefore overestimates severity of the modelled cohort). The mean age of patients at baseline was 39.3 years with the female to male ratio 2.98:1. The baseline EDSS distribution is presented in GMS Figure B7.3.1, page 243. However, the EDSS data does not seem to match the published literature (Pickin[32] or Boggild[34]), so it is unclear where this has been sourced and whether it is representative of the UK population. (Clarification requested; source of data stated in clarifications as Pickin[32] so discrepancies exist between reported source and data used in the model; however, these are not a major concern.) As alemtuzumab is expected to be used in treatment naïve as well as treatment experienced patients the ERG considers the less severe distribution of EDSS states from the trial population to be more appropriate for the economic analysis.

The economic evaluation considers two subgroups within active RRMS, HA despite treatment with beta interferon and RES, identified by clinical experts and NICE as appropriate potential subpopulations in which alemtuzumab could be used. Two subgroups not considered in the economic evaluation are treatment naïve and treatment experienced patients.

**Interventions and comparators**

The recommended dose of alemtuzumab is 12 mg/day administered by intravenous infusion for 2 treatment courses (initial one for five consecutive days, second 12 months later for three consecutive days). For the cost effectiveness analysis the comparators are beta-interferons, glatiramer acetate, and natalizumab (for treatment naïve or previously treated patients with RES) and fingolimod (for patients with HA who have received treatment) as specified in the NICE scope. All comparators are current treatment options used in the NHS for MS. In line with Association of British Neurologists guidelines the costs and benefits of treatments include a stopping rule such that when patients reach EDSS 7 or convert to SPMS they receive the same benefits as a best supportive care (BSC) patient.

**Clinical Effectiveness**

The clinical effectiveness parameters used in the model include disease progression (as 3-month SAD), relative risk of relapse (as ARR), withdrawal probability, and relative risk of relapse leading to hospitalisation (GMS Table B7.3.24, page 271). The key clinical effectiveness parameters used in the model are shown in Table 28. The main source for the treatment effects used in the model is the manufacturer's post-2000 MTC and these effects were applied to the

natural history of progression and relapse. See Table 4 for ERG critique of the MTC. Overall, there are no particular methodological issues relating to the MTC which may bias results although results for treatment naïve and treatment experienced patients are combined which may not be appropriate.

**Table 28: Key clinical effectiveness parameters used in the model**

| Treatment | 3-month SAD HR | Relative ARR (from : Base case MTC: Post-2000, 80% RRMS MTC) | Relative proportion of relapses leading to hospitalisation |
|---|---|---|---|
| Alemtuzumab | [redacted] | [redacted] | 0.225 |
| IM IFNβ-1a | 0.91 | 0.78 | 0.495 |
| IFNβ-1b | 1.21 | 0.68 | 0.495 |
| SC IFNβ-1a 44µg | 0.79 | 0.62 | 0.495 |
| SC IFNβ-1a 22µg | [redacted] | 0.62[1] | 0.495 |
| Glatiramer acetate | 0.93 | 0.64 | 0.495 |
| Fingolimod | 0.75 | 0.46 | 0.600 |
| Natalizumab | 0.58 | 0.31 | 0.600 |

[1] Efficacy of 22 µg assumed to equal 44 µg.

*Disease progression*

The natural history was modelled using the London Ontario dataset[29] as this was considered the most appropriate natural history dataset for active RRMS and SPMS patients not receiving DMT. As no data for EDSS 0 were available these transition probabilities were obtained from the placebo arms of the TOWER and TEMSO trials of teriflunomide. Data for EDSS 9 was also not available from the London Ontario dataset and a 100% conversion rate for RRMS patients in EDSS 9 to SPMS was assumed. The natural history transition matrix of active RRMS patients is reported in GMS Table B7.3.4 (GMS page 249) and that of SPMS patients in GMS Table B7.3.5 (GMS page 250). The natural history transition matrix is calculated using the active RRMS and SPMS natural history data with the probability of conversion from RRMS to SPMS (GMS Table B7.3.7, page 252).

The probability of converting from RRMS to SPMS is calculated from HRs derived from time to event data and survival analysis and the annual probability of conversion to SPMS from RRMS is given in GMS Table B7.3.6 (GMS page 251). The current method for deriving probabilities of

conversion from RRMS to SPMS was also used in a recent NICE technology appraisal for teriflunomide and was criticised in the ERG report (York).[35] This report states that the manufacturer's estimates over predict the rate of conversion for a patient starting in EDSS 1 such that the median time to conversion from EDSS1 to SPMS is 10 - 11 years, which is considerably less than the 15 year conversion seen in the literature. In order to assess the impact of conversion rate on cost-effectiveness, the ERG has undertaken an additional analysis using the conversion rates estimated by the York ERG.

*Treatment effect on disease progression – natural history method*

For the base case, HRs for 3-month SAD obtained from the MTC were applied to the natural history matrix to derive treatment transition matrices. Although the London Ontario dataset has been used in previous technology appraisals (e.g. Chilcott and colleagues[20]) and has been used in the current model for consistency, the ERG has concerns about this approach. There are a number of limitations of the natural history method: it does not allow for regression in EDSS state (patients can only progress in EDSS or to remain in their existing EDSS for either RRMS or SPMS patients); it uses HRs derived from a MTC which combines results for treatment naïve and treatment experienced patients; it applies HRs to a natural history dataset which may be out-of date and not representative of patients likely to receive alemtuzumab (collected 1972 to 1984, n=345 patients); the natural history dataset did not provide data for EDSS 0 and 9; and EDSS progression in untreated MS patients may be much slower than previously estimated (e.g. twice as long as the 15 years estimated in the London Ontario cohort[26]).

*Treatment effect on disease progression – direct comparison method*

A sensitivity analysis was undertaken which used data from the CARE-MS clinical trials to derive treatment transition matrices. The transition matrices for this direct analysis were estimated from the treatment arms in the alemtuzumab phase III trials and reflect patients' SAD as observed whilst receiving alemtuzumab or SC IFNβ-1a 44 µg and are shown in GMS Tables B7.3.21 (page 267) and B7.3.22 (page 266) respectively. Whilst using trial data allows for patients to improve their disability, there are other limitations of this method particularly the use of various different sources of data and the fact that 6 month SAD data is available from the CARE MS trials which could have been used instead of 3 month as it is a more robust measure. The direct comparison method did not provide individual patient data for transitions to EDSS 9 and from EDSS 8 to 9 (which were supplied from the natural history progression data from London Ontario[29]); where numbers of transitions are fewer than 15 these are also derived from

London Ontario; BSC is used when patients withdraw from treatment and this uses the London Ontario dataset transition probabilities; the trial data is short-term and long-term open label extensions of RCTs have not demonstrated significant benefit of DMTs in preventing disability progression as measured by EDSS.[36]

Comparison of the two methods of estimating SAD is presented in GMS Table B7.3.23 (page 268).

*Relapse rate*

To evaluate the effect of treatment on relapse rate, the ARR derived from the MTC was applied to the baseline natural history relapse rate. As no relapse data were available from the London Ontario dataset, the natural history relapse rates were sourced from alternative literature sources. No details are given about the methods used for identifying the literature. Two studies were included, one by Held and colleagues[37] was used in the base case since it is more recent and Patzold and colleagues[38] as per previous submissions (TA 127[28] and TA 254[27]) is considered in a sensitivity analysis.

The natural history relapse rates by MS classification and EDSS state were calculated in a two-step process. Firstly mean ARRs by years since diagnosis were derived from Held and colleagues[37] and Patzold and colleagues.[38] Then these rates were applied to the UK MS Survey number of patients in each EDSS state by number of years since diagnosis to give the number of relapses per EDSS state per year since diagnosis (GMS Table B7.3.12, page 255). The ARR derived from the MTC was then applied to the natural history relapse rates to give the relative risk of relapse rates compared to placebo due to treatment (GMS Table B7.3.13).

*Relapse severity leading to hospitalisation*

Relapse severity is also included in the model and data sourced from the systematic literature review of natural history although no details are given of methods used. One study was identified in which 20% of relapses lead to hospitalisation (Dee and colleagues[39]), which is stated to be representative of the UK population. This proportion is applied to the treatment effect on ARR to give the number of relapses leading to hospitalisation. Assumptions and sources of data are provided in GMS Table B7.3.14 (page 257). These assumptions seem reasonable although clinical advice to the ERG suggests that increasingly relapses are treated

Superseded – see Erratum

in the community and the model may be overestimating the proportion of relapses that are treated in hospital.

*Withdrawal from treatment*

The impact of withdrawing from treatment was assumed to have no effect (due to the sustained effect of alemtuzumab long after treatment) and was therefore not modelled explicitly. However, for all comparators patients may withdraw from treatment at the start of each cycle. The method used to calculate the probability of treatment withdrawal is based on the 2 year withdrawal probability for teriflunomide combined with the results of the MTC to give annual probabilities of withdrawal for all treatments (for consistency with the GMS of teriflunomide to NICE May 2013). It is assumed in the base case that the probability of withdrawal is reduced by 50% after year 2, as it is anticipated that a patient is likely to be more tolerant to adverse events and that all-cause discontinuation would decrease. Annual all-cause rates of withdrawal used in the model are presented in GMS Table B7.3.2. Clinical advice to the ERG suggests that tolerance is improved over time but there is the potential for new/aggravation of side effects with additional treatment.

Treatment related adverse events at an incidence of ≥4% compared with placebo or if no placebo data available, >5% in the treatment arms, were included in the model. These were arbitrary thresholds which were expected to capture the most common adverse events that occur with DMTs. The sources of data for each DMT were placebo controlled studies that had been included in the MTC. Probabilities of adverse events for alemtuzumab for the model were derived from Pool C of the Integrated Summary of Safety (Genzyme 2012). Methods for calculating probabilities for the first year and subsequent years are discussed and seem reasonable but these data has not been checked as the ERG did not have access to it. It was assumed in the GMS that adverse events associated with alemtuzumab may persist up to EDSS 7 in RRMS (in line with the assumption that treatment effect of alemtuzumab persists up to EDSS 7 in RRMS regardless of dosage); once patients reach EDSS 7 or SPMS they are assumed to receive no adverse events for alemtuzumab. Adverse event probabilities used in the model are shown in GMS Table 7.3.17 page 261.

**Patient outcomes**

The cost-effectiveness model incorporated the health-related quality of life (HRQoL) impact of the different treatments into QALYs. HRQoL data used in the model accounted for EDSS/SPMS level, relapse occurrence, treatment-related adverse events and carer disutility.

The GMS states that HRQoL data from the CARE-MS trials[3;4] was not available for implementation into the model (GMS page 283). Therefore a systematic review of the literature was conducted to obtain all relevant HRQoL studies in MS. The inclusion criteria for the HRQoL review are shown in GMS Table C10.12.3 (GMS Appendix 10.12, page 451). Studies were included if they reported utility in MS or disutility of relapse, adverse events or to carers for UK adults with either RRMS or progressive MS. Studies in PPMS were excluded.

Of the ten relevant studies which met the inclusion/exclusion criteria two were implemented in the model. No justification is given as to why these two studies were selected (Orme and colleauges[40] and Gani and colleagues[22]) nor any details presented about the remaining eight studies. Of these two studies, one is a cross sectional study of people with all three types of MS using a postal questionnaire completed by patients or carers of patients identified through the UK MS Trust (UK MS Survey, Orme and colleagues[40]) and the other study contained HRQoL data using patients from the Affirm RCT and the UK MS Survey (Gani and colleagues[22]).

Previous use of the utility data from the study by Orme and colleagues[40] has been heavily criticised in the PenTAG ERG report on natalizumab (TA127),[41] highlighting the low response rates (20%), selection bias, the unrepresentative population and self-reported severity estimates. The York ERG report on teriflunomide also considered this data problematic and recommended using trial-based utility values.[35] However, another systematic review of utilities in MS[42] identified 16 studies reporting utilities associated with health states in MS as measured by EDSS, 3 of which were UK studies. Whilst the utilities ranged across EDSS categories, results showed that utilities from the Orme and colleagues study[40] are similar to the other studies and in the absence of better evidence it seems reasonable to use this data.

Health state utilities were applied for each EDSS state in the model and based on a published regression of quality of life responses from a survey of patients and carers of patients with MS. EQ-5D utility scoring system was applied with respondent domain scores converted to a single utility weight using the UK value set (Orme and colleagues[40]). EDSS utilities are shown in GMS Table B7.4.2 (page 288) and Table B7.4.5 (page 294). It is assumed in the study (and hence in the model) that patients who have converted to SPMS have a fixed utility decrement of 0.045 over the corresponding RRMS EDSS state utility values (sourced Orme and colleagues[40]). EDSS utilities used in the model are shown in Table 29.

**Table 29: EDSS utilities used in the model (from GMS Table B7.4.2)**

| EDSS state | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| RRMS | 0.870 | 0.799 | 0.705 | 0.574 | 0.610 | 0.518 | 0.460 | 0.297 | -0.049 | -0.195 |
| SPMS | 0.825 | 0.754 | 0.660 | 0.529 | 0.565 | 0.473 | 0.415 | 0.252 | -0.094 | -0.240 |

Utility loss for recent relapse (-0.071) was also derived from Orme and colleagues[40] Disutility of relapses leading to hospitalisation is higher than relapses that do not lead to hospitalisation.[39] This increased disutility was applied to the average UK disutility of relapse to derive a disutility of relapses leading to hospitalisation. Thus the modelled utility loss per relapse event was -0.071 for events not leading to hospitalisation and -0.2356 for events leading to hospitalisation. The average duration of relapse was sourced from Gani and colleagues[22] (1.51 months) and combined with the disutility during a relapse to give the disutility per relapse.

Disutilities of adverse events are presented in GMS Table 7.4.4 (page 290). The total disutility of an event occurring is calculated using the disutility multiplied by the expected duration of the event. A range of adverse events using different sources and assumptions are given which seem comprehensive. These have not been checked in detail by the ERG due to time constraints imposed by the size of the GMS but are likely to have limited impact on model results.

Disutility for carers has been included in the model using a method developed by Gani and colleagues[22] which assumed that disutility had a maximum value of 0.14. This was based on the value accepted by NICE in an assessment of treatments for Alzheimer's Disease.[43] Disutility for MS carers by patient EDSS score was calculated by multiplying the maximum disutility of 0.14 by the percentage of time spent caring, which occurred at EDSS 8.5 - 9.5. This provided an index of disutilities of 0.00 at EDSS to 0.14 at EDSS 8.5 - 9.5. The percentage of time spent caring for a person with MS was obtained from the UK MS Survey by EDSS score. Carer disutilities used in the model are shown in GMS Table B7.4.3 (page 289).

One-way sensitivity analyses changing the EQ-5D utilities by EDSS score and MS classification using 95% confidence intervals and disutilities by plus or minus 10% did not impact on results. The ERG has undertaken a scenario analysis using a different source of utilities to determine their differential effect on cost-effectiveness findings.

Superseded – see Erratum

**Resource use**

The cost analysis was conducted from an NHS and PSS perspective with 2012 used as the costing year. The resource use categories used within the model were: drug acquisition, administration and monitoring costs. The GMS reports a systematic review to obtain all relevant cost and resource use studies in MS. Details are provided in GMS Appendix 13 (page 453). Eighteen studies met the review inclusion/exclusion criteria of which three were implemented in the model (GMS Table B7.5.6, page 308). No justification is given as to why these studies were selected: Tyas and colleagues[44] is the source for health states and associated costs in the base case, with Karampampa and colleagues[45] used for sensitivity analysis values; Dee and colleagues[39] is used for resource use and costs associated with relapses.

*Drug acquisition*

The treatment regimen for alemtuzumab was based on that in the CARE-MS trials,[3;4] using an initial course of 12 mg/day (IV) for 5 consecutive days and subsequently 12 mg/day for 3 consecutive days administered 12 months after initial treatment.

The GMS states that the majority of patients will only require two courses of treatment as the effect of alemtuzumab is assumed to persist over the long-term and so the continued benefit of alemtuzumab is modelled such that patients receive the full efficacy of alemtuzumab when they do not receive a course of treatment. As annual acquisition costs for alemtuzumab are therefore dependent on the proportion of patients who have subsequent course of treatment, in order to avoid underestimation of acquisition costs the base case assumes that [redacted] of patients receive a subsequent dose in year 3, based on the CAMMS 223 extension study with [redacted] in years 4 and 5. Beyond year 5, a [redacted] annual rate of retreatment is assumed. The ERG assess the effect of varying this assumption on the model results in section 4.3.

The annual drug acquisition for the comparator DMTs is the same for all patients on treatment. Source for beta-interferons and glatiramer acetate were from the Department of Health (DH) Health Service Circular 2002 and BNF 2013 for fingolimod and natalizumab.[46-48]

*On treatment monitoring and management*

The model also included concomitant medications as recommended in the alemtuzumab SmPC and in the clinical trials. Patients were pre-treated with 1 g methylprednisolone (to reduce

allergic reactions) and acyclovir 200 mg twice daily for 28 days (as prophylaxis for herpes infection). In addition, pre-treatment with chlorphenamine (antihistamine) and paracetamol (antipyretic) is also included. Resources for beta-interferons and glatiramer acetate include initial one-off training by a clinical nurse specialist for self-injection; continuous ECG and BP monitoring for fingolimod following first dose; methylprednisolone administration with IV infusion of natalizumab 13 times per year (GMS Table B7.5.3, page 303).

Monitoring resource use associated with treatments was derived from SmPCs and the NICE costing template (2012) and is presented in GMS Table B7.5.4 (page 305). Neurologists were consulted to assess whether these reflected clinical practice. Responses varied in terms of number of liver function tests, full blood counts and neurology visits and in each case a conservative approach was taken and the NICE costing template used as source of data (justification being that an underestimate would apply to all DMTs and not significantly change to ICER). For alemtuzumab it is assumed that monitoring persists after dosing and/or discontinuation; this is a conservative assumption as the SmPC states that patients would only receive 4 years of monitoring following their last course of treatment.

The included resource use appears relevant and comprehensive.

**Costs**

The acquisition cost of alemtuzumab from Genzyme is indicative and to be confirmed with DH. Unit acquisition costs of alemtuzumab (indicative price) are £7,045 per vial, first 5 vials £35,225, second 3 vials £21,135 (GMS Table B7.5.7.). Acquisition costs for beta-interferons and glatiramer acetate were obtained from DH Health Service Circular 2002 and from the BNF 2013 for fingolimod and natalizumab.  Administration costs were obtained from the Personal Social Services Research Unit (PSSRU) for beta-interferons and glatiramer acetate and NHS Reference Costs 2011-12 for alemtuzumab, fingolimod and natalizumab. Monitoring costs come from a range of sources including NHS Reference costs 2011-12 and Payment by Results tariffs 2012-13. Costs are shown in Table 30.

**Table 30: Annual costs associated with the technologies in the economic model (from GMS Table B7.5.7)**

| Treatment | Acquisition cost, £ | | Administration cost, £ | | Monitoring cost, £ | | Total, £ | |
|---|---|---|---|---|---|---|---|---|
| | Year 1 | Year 2+ | Year 1 | Year 2+ | Year 1 | Year 2+ | Year 1 | Year 2+ |
| Alemtuzumab | 35,225.00 | 21,135.00 | 2,438 | 1,487 | 443.00 | 274.00 | 13,936.12 | 13,897.24 |
| SC IFNβ-1a 44µg | 8,942.00 | 8,942.00 | 174.00 | 0 | 355.28 | 346.64 | 9,471.28 | 9,288.64 |
| SC IFNβ-1a 22µg | 7,513.00 | 7,513.00 | 174.00 | 0 | 355.28 | 346.64 | 8,042.28 | 7,859.64 |
| IM IFNβ-1a | 8,502.00 | 8,502.00 | 174.00 | 0 | 355.28 | 346.64 | 9,031.28 | 8,848.64 |
| IFNβ-1b | 7,259.00 | 7,259.00 | 174.00 | 0 | 355.28 | 346.64 | 7,788.28 | 7,605.64 |
| Glatiramer acetate | 5,823.00 | 5,823.00 | 174.00 | 0 | 338.00 | 338.00 | 6,335.00 | 6,161.00 |
| Fingolimod | 19,162.50 | 19,162.50 | 474.00 | 0 | 641.74 | 346.64 | 20,278.24 | 20,001.14 |
| Natalizumab | 14,730.00 | 14,730.00 | 6162.00 | 6162.00 | 493.46 | 493.46 | 21,385.46 | 21,385.46 |

*Health state resource use*

Health states and associated costs in the economic model are taken from two of the studies identified by the systematic review for resources and costs (Tyas and colleagues[44] for the base case, Karampampa and colleagues[45] for sensitivity analysis). These provide direct medical and other direct costs by EDSS which apply to both RRMS and SPMS and are presented in GMS Table B7.5.8, (page 313). As can be seen in Table 31, there were large differences in health state costs between the two sources of data. As details of the constituents of the direct costs are not given, it is not clear what is included and it is difficult to assess if the approach is consistent with the NHS and PSS perspective, and there is some uncertainty around the correct value for health state costs. The most severe states of disability incur the greatest costs which is plausible as PSS social care costs are likely to be considerable for the more severe health states.

**Table 31: List of health states and associated costs in the economic model (from GMS Table B7.5.8)**

| Health states (applies to RRMS and SPMS) | Base case value (Tyas *et al.* 2007[44]) | Sensitivity analysis value (Karampampa *et al*. 2012[45]) |
|---|---|---|
| EDSS 0 | £5670.77 | £3579.456 |
| EDSS 1 | £5979.77 | £3579.456 |
| EDSS 2 | £7134.19 | £3579.456 |
| EDSS 3 | £10880.51 | £3579.456 |
| EDSS 4 | £7755.90 | £14171.7 |

| | | |
|---|---|---|
| EDSS 5 | £11545.48 | £14171.7 |
| EDSS 6 | £12837.10 | £14171.7 |
| EDSS 7 | £24356.62 | £49661.93 |
| EDSS 8 | £34616.65 | £49661.93 |
| EDSS 9 | £32619.28 | £49661.93 |
| EDSS 10 | £0 | £0 |

*Cost of relapses*

The model differentiates the severity of relapses by relapses leading to hospitalisation and those that do not. The resource use of having had a relapse leading to hospitalisation was estimated from Irish neurology centres giving a length of stay of 10.71 days; it was assumed that relapses not leading to hospitalisation required 5 days for steroid treatment.[39] UK NHS reference costs were used with the resource use to estimate costs of relapse by hospitalisation or not. These are shown in Table 32.

**Table 32: Costs associated with relapses included in the economic model**

| Category | Base case cost (Dee, 2012) |
|---|---|
| Relapse not leading to hospitalisation | £ 844.65 |
| Relapse leading to hospitalisation | £ 6,164.46 |

Overall, cost and resource use parameters were varied by 10% in the sensitivity analyses and are shown not to have a large impact on the cost effectiveness results.

**Consistency/ Model validation**

**Internal consistency**

The electronic model is coded in Microsoft Excel and is fully executable. The model is well presented and documented and user friendly.

The GMS states that quality assurance of the model included two independent health economists involved in the design and build, as well as clinical expertise ratifying plausibility of results. The ERG have not undertaken a comprehensive check of all cells in the model, rather random checking of the model has been done for some of the key equations in the model. Changing the parameter values produced intuitive results and from random checking the 'wiring' of the model appears to be accurate. The ERG was able to replicate the results presented in the GMS and the deterministic sensitivity analyses. The ERG views the model as a reasonable approach to modelling the cost effectiveness of MS.

**External consistency**

The model outputs have been compared to published literature to validate outputs (GMS Section 7.7.1). The model results are compared for alemtuzumab versus IFN- β1a (Rebif) 44µg between clinical and model outputs. The results were compared at the end of year 2. There was reasonable agreement, with generally more relapses in the model than the trials. The GMS states this is likely to be a consequence of relapses being calculated based on EDSS state and not using the clinical trial data directly. The quality of life of the model was also lower. The GMS states that this is due to the fact that patients are in a worse EDSS state when starting the model and are not able to regress which therefore may accelerate their decline leading to poorer quality of life.

The ERG notes that there is no longer term validation beyond 2 years, and therefore there is some uncertainty around the validity of the longer term outcomes. In addition, the analysis has not used the same baseline characteristics as the trials, as the model population was based upon the UK RSS, as so the ERG considers the validation to be flawed.

**Assessment of Uncertainty**

The manufacturer has assessed uncertainty within the model by conducting sensitivity and scenario analysis for structural assumptions and parameter input values. The GMS contains deterministic and probabilistic sensitivity analyses and scenario analyses.

**One-way sensitivity analyses**

Deterministic sensitivity analyses were run for the base case of alemtuzumab versus SC IFNβ-1a 44 µg, fingolimod and natalizumab. For active RRMS all of the parameters were varied in the one-way sensitivity analysis. For HA disease despite treatment with a beta interferon, and RES, only the top 5 most sensitive parameters were analysed. The GMS justified the use of SC IFNβ-1a 44 µg on the grounds that it was the most efficacious treatment of the beta-interferons and glatiramer acetate according to the MTC, in term of reduction in SAD and relapse (Table 28). Furthermore, the CARE-MS trials[3;4] provided a direct comparison of alemtuzumab versus SC IFNβ-1a 44 µg. The ERG considers that this is a reasonable and pragmatic approach.

The input parameters that the model is most sensitive to are shown in Table 33. These are the HR on SAD for alemtuzumab and SC IFNβ-1a 44 µg, the disease costs and the withdrawal rate for SC IFNβ-1a 44 µg. The model results are very sensitive to the treatment effect on SAD, with

cost effectiveness varying between alemtuzumab dominating SC IFNβ-1a 44µg (for upper confidence interval of HR SAD of 0.3) to not being cost effective (> £1 million per QALY for lower confidence interval of HR SAD of 0.9). The model results are insensitive to other changes to the model input parameters.

**Table 33: One way sensitivity analysis results of alemtuzumab versus SC IFNβ-1a 44 µg in active RRMS (GMS Table B7.7.15)**

| | | **Maximum** | | | **Minimum** | | |
|---|---|---|---|---|---|---|---|
| | **Variation** | **Inc Cost £** | **Inc QALY** | **ICER £** | **Inc Cost £** | **Inc QALY** | **ICER £** |
| Base case | | 9,993 | 1.183 | 8,445 | | | |
| Alemtuzumab HR on sustained accumulation of disability | 95% CI | -11,391 | 2.152 | Dominates | 35,696 | 0.030 | 1,200,973 |
| SC IFNβ-1a 44µg HR on sustained accumulation of disability | 95% CI | 18,642 | 0.756 | 24,668 | -325 | 1.695 | Dominates |
| Disease costs | +/- 10% | 12,898 | 1.183 | 10,900 | 7,088 | 1.183 | 5,990 |
| SC IFNβ-1a 44µg withdrawal rate | 95% CI | 12,946 | 1.220 | 10,613 | 6,356 | 1.139 | 5,580 |

HR, Hazard ratio.

The sensitivity of the incremental results to the clinical parameters sourced from the MTC was tested in the GMS for the full sensitivity analysis MTCs of studies from all-years and 100% RRMS rather than the base case of post-2000 80% RRMS (GMS Tables B7.7.18 - 20). For these analyses, the cost effectiveness of alemtuzumab versus glatiramer acetate was £9,982 / QALY for all years, 80% RRMS; £27,434 / QALY for all years, 100% RRMS and £10,822 / QALY for post-2000, 100% RRMS. All other treatments were dominated by either alemtuzumab or glatiramer acetate for all analyses.

The manufacturer performed structural deterministic sensitivity analysis (DSA) for changes to the discount rate, the waning effect of treatment, duration of autoimmune disease, and assumptions around the treatment effect on relapses (GMS Table B7.7.21, page 351). Selective analyses are shown in Table 34. They also presented sensitivity analyses using different data sources for the baseline characteristics of the MS patients, the natural history costs, and the natural history transition probabilities, and alternative long term dosage of alemtuzumab. The GMS analyses showed that results were sensitive to long term waning of efficacy, natural history costs and assumptions regarding dosage of retreatment with alemtuzumab.

The justification for the structural sensitivity analyses have been provided in Table B7.3.25. The ERG considers that the GMS has performed appropriate structural sensitivity analyses and the rationale for these analyses is reasonable. The ERG notes that some of these structural sensitivity analyses have been recommended in previous MS technology appraisals.[27] The ERG considers that the sensitivity analyses completed are reasonably comprehensive, although there are no sensitivity analyses that vary the disease progression for best supportive care, or the progression from RRMS to SPMS. The ERG has assessed the impact of varying these parameters on the model results in section 4.3.

**Table 34: Deterministic parameter sensitivity analysis of alemtuzumab versus SC IFNβ-1a 44µg in active RRMS (GMS Table B7.7.21)**

| Scenario | Alemtuzumab total costs (£) | Alemtuzumab total QALYs | Inc costs (£) | Inc QALYs | ICER (£) Inc (QALYs) |
|---|---|---|---|---|---|
| Base case | 499,347 | 4.034 | 9,993 | 1.183 | 8,445 |
| Baseline characteristics from CARE-MS studies | 480,425 | 6.417 | 1,412 | 1.624 | 869 |
| Natural history costs from Karampampa *et al.* 2012[45] | 653,402 | 4.034 | -14,334 | 1.183 | Dominates |
| Natural history transition probabilities for All RRMS | 496,816 | 4.143 | 10,147 | 1.180 | 8,597 |
| Long-term treatment effect 25% waning after year 5 for all treatments | 503,798 | 3.845 | 14,095 | 1.010 | 13,956 |
| Long-term treatment effect 50% waning after year 5 for all treatments | 507,638 | 3.683 | 17,602 | 0.863 | 20,388 |
| No waning of discontinuation | 499,347 | 4.034 | 12,926 | 1.218 | 10,617 |
| No treatment effect on relapses or proportion leading to hospitalisation | 508,523 | 3.975 | 16,546 | 1.140 | 14,517 |
| Using direct comparison method: transition probabilities derived directly from pooled CARE-MS I[2] and CARE-MS II[3] | 377,329 | 10.236 | -72,366 | 5.410 | Dominates |

**Scenario Analysis**

Separate analyses were completed for the subgroups of patients with HA RRMS despite treatment with a beta interferon and RES RRMS. The deterministic results are shown in GMS Tables B7.9.4 and B7.9.6. The input parameters for these analyses were from a MTC conducted for these subgroups of patients. For HA RRMS despite treatment with a beta interferon, alemtuzmab dominated fingolimod (assuming a patient access scheme price of £13,000), and for RES RRMS, alemtuzumab dominated natalizumab. The manufacturer also ran probabilistic and deterministic sensitivity analyses for these subgroup analyses.

**Probabilistic Sensitivity Analysis**

The PSA uses 10,000 iterations and runs in approximately 5 hours. The results are presented as incremental probabilistic base case results versus all comparator (GMS Table 7.7.9). The GMS presents probabilistic results as the base case results, and state that this had been recommended by the ERG reviewing the fingolimod STA submission.[43] The probabilistic results are consistent with the deterministic results (ICERs of £7,017 / QALY [probabilistic] versus £8,924 [deterministic]) for alemtuzumab versus glatiramer acetate.

Variables and their distributions included in the PSA are reported in GMS Table 7.6.4 (page 323). Parameters for these distributions are not provided in the GMS but they are supplied in the economic model. The PSA includes most model input parameters, except for the natural history transition probabilities and the drug acquisition costs. The GMS states that they omitted the natural history probabilities from the PSA because there are small sample sizes in high RRMS EDSS states and low SPMS EDSS which result in unrealistic stochastic transition probabilities. As the natural history probabilities are a large source of uncertainty in the model, the ERG considers that they should be included within the PSA and therefore the uncertainty has not been fully explored. The ERG considers that in general the probability distributions are correctly applied, although the gamma distribution has been used for the utility and disutility whereas the beta distribution is more usual.

The PSA was performed for alemtuzumab versus all comparators in active RRMS. A multi-CEAC is presented in Figure 2 (GMS Figure B7.7.7, page 350). Alemtuzumab has the highest probability of being cost effective at a willingness to pay threshold above £3000 / QALY. At a willingness to pay threshold of £20,000 / QALY, the probability of alemtuzumab being cost effective is 35%.

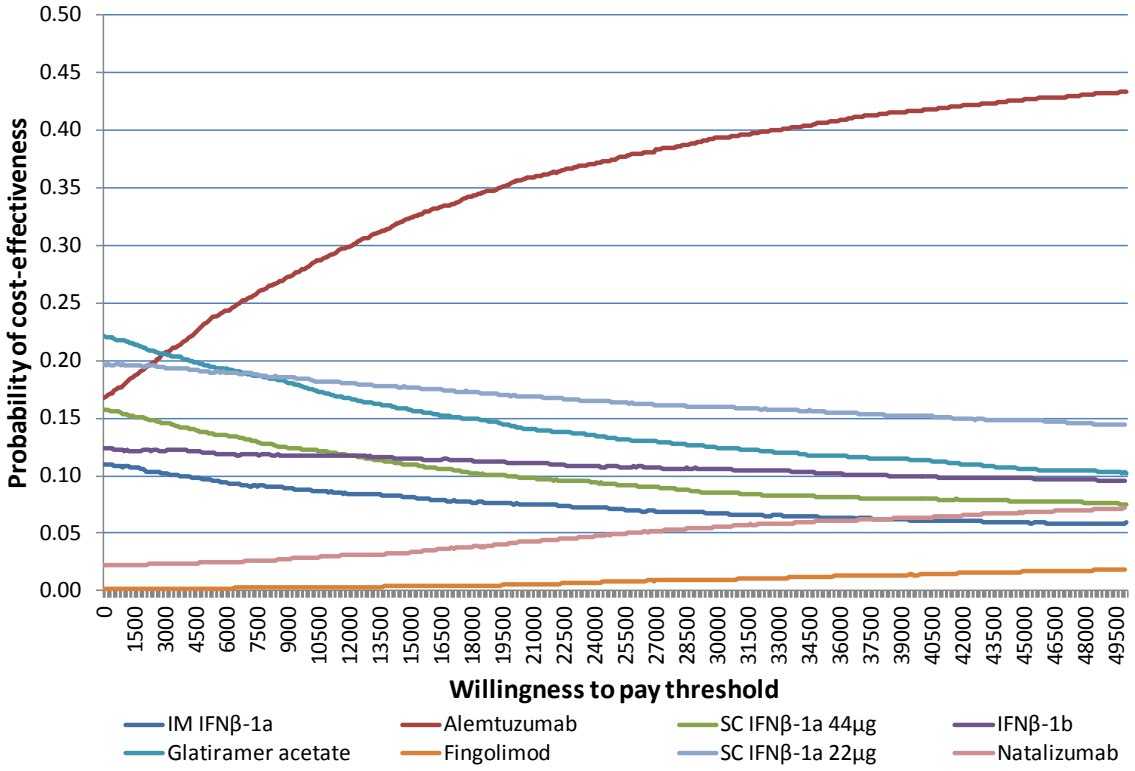


**Figure 2: Multi-CEAC of alemtuzumab versus all beta-interferons, glatiramer acetate, fingolimod and natalizumab in active RRMS (GMS Figure B7.7.7)**

**Comment on validity of results with reference to methodology used**

The structure adopted for the economic model is reasonable and consistent with current clinical understanding of MS and previous economic evaluations of treatments for MS. The methods of analysis are appropriate and conform to NICE methodological guidelines. The parameters used for the model are generally appropriate.

Previous NICE appraisals of beta-interferons and glatiramer aceteate (TA 32),[49] natalizumab (TA 127)[28] and fingolimod (TA 254)[27] have estimated the costs and QALYs for patients with MS. The York ERG has analysed these appraisals in order to provide external validity for a recent NICE appraisal for teriflunomide.[35] They concluded that the model was underestimating QALYs. They concluded that a better approach would be to use the trial distribution of initial EDSS, as this was more reflective of first-line patients. They also concluded that the conversion rate for patients to transition from RRMS to SPMS was too high.

As the modelling of teriflunomide is similar to that of alemtuzumab, the York ERG conclusions regarding validity are also relevant for this appraisal. The ERG agrees with these criticisms and has re-run the model using these assumptions. These results are shown in section 4.3.

The York ERG also recommends that the trial estimates of natural history should be used in the model as this includes improvements in progression and was more reflective of first-line patients, and that patient costs should not include non-medical costs from Tyas and colleagues.[44] These recommendations are problematic, because the natural history transition probabilities from the trial would be based on only a short time period, and the patient costs in Tyas and colleagues[44] are not clearly defined and so may contain relevant PSS costs. Therefore we vary some of these assumptions in sensitivity analyses in Section 4.3.

The York ERG also considered that best supportive care should be included within the NICE scope. We have included this for information in section 4.3.

### 4.3 Additional work undertaken by the ERG

This section details the ERG's further exploration of the issues and uncertainties raised in the review and critique of the GMS cost effectiveness analyses. The ERG presents a preferred base case, with an alternative patient population and different progression rate from RRMS to SPMS. A series of sensitivity analyses are then run for this new base case.

The ERG expressed concern on the population used for the analyses. The modelled population represents all RRMS patients, although the population of interest is those patients that would otherwise use alemtuzumab or a comparator treatment. The GMS model provides four alternative population options: UK RSS, TEMSO and TOWER, Pooled CARE-MS trials, and AFFIRM. Of these, the ERG considers that the pooled trial population from the CARE-MS trials[3;4] is the most relevant population. The ERG compared alemtuzumab versus SC IFNβ-1a 44 µg. SC IFNβ-1a 44 µg was used as the comparator in the GMS for the sensitivity analyses because it was the direct comparator in the clinical trials and it was the most efficacious in the MTC. For this reason, we have used it as a comparator in the ERG analyses. For the analysis with the CARE-MS trials[3;4] patient characteristics (Table 35), the ICER reduces from £8445 (GMS base case) to £2865 per QALY gained.

The York ERG analysed the conversion rate from RRMS to SPMS model.[35] They concluded that the conversion rate used for the teriflunomide analysis was too high and recommended a lower rate (roughly half the modelled rate). As the GMS used the same conversion rate as for terflunomide, we considered that this approach was also appropriate for this appraisal. Table 35 shows the effect of reducing the conversion rate to that recommended by the York ERG, i.e. the ICER reduces to £3100 per QALY gained for alemtuzumab versus SC IFNβ-1a 44 µg.

The ERG's preferred approach uses both the population from the CARE-MS trials[3;4] and the reduced conversion rate from RRMS to SPMS. The effect of these changes is shown in Table 35, where the cost effectiveness of alemtuzumab becomes more favourable and now dominates SC IFNβ-1a 44µg, i.e. cheaper and more effective.

**Table 35: Incremental deterministic base-case results of alemtuzumab versus SC IFNβ-1a 44µg**

| | SC IFNβ-1a 44µg | | Alemtuzumab | | | | |
|---|---|---|---|---|---|---|---|
| **Technologies** | **Total costs (£)** | **Total QALYs** | **Total costs (£)** | **Total QALYs** | **Incr. costs (£)** | **Incr. QALYs** | **ICER (£) inc. (QALYs)** |
| Base case | 489,354 | 2.850 | 499,347 | 4.034 | 9,993 | 1.183 | 8,445 |
| Trial characteristics | 439,732 | 4.894 | 444,226 | 6.460 | 4494 | 1.566 | 2,869 |
| Reduced progression RRMS to SPMS | 480,755 | 3.083 | 485,379 | 4.575 | 4624 | 1.492 | 3,100 |
| Trial characteristics and reduced progression RRMS to SPMS | 431,896 | 5.205 | 430,241 | 7.147 | -1655 | 1.942 | -852[a] |

[a] Alemtuzumab dominates SC IFNβ-1a 44 µg, ie is cheaper and more effective.

Table 36 shows the ERG's preferred base case for all comparative treatments. Although best supportive care was not in the NICE scope, this comparator was available in the GMS model and we have provided this for contextual information. In this case alemtuzumab is shown to dominate all comparative treatments and be cost effective compared to best supportive care with an ICER of £9907 per QALY gained.

**Table 36: ERG Incremental deterministic base-case results of alemtuzumab versus SC IFNβ-1a 44 µg, BSC and glatiramer acetate (with trial characteristics and reduced progression RRMS to SPMS)**

| Technologies | Total costs (£) | Total QALYs | Inc costs (£) vs BSC | Inc QALYs vs BSC | ICER (£) inc (QALYs) |
|---|---|---|---|---|---|
| Best supportive care | 408,040 | 4.906 | - | - | - |
| Alemtuzumab | 430,241 | 7.147 | 22,201 | 2.241 | 9907 |
| Glatiramer acetate | 430,635 | 5.065 | 22,595 | 0.159 | Dominated |
| SC IFNβ-1a 44µg | 431,896 | 5.205 | 23,856 | 0.299 | Dominated |
| IM IFN | 440,185 | 5.089 | 32,145 | 0.183 | Dominated |
| IFN | 453,837 | 4.431 | 45,797 | -0.475 | Dominated |
| Fingolmod | 492,053 | 5.539 | 84,013 | 0.633 | Dominated |
| Natalizumab | 493,466 | 5.962 | 85,426 | 1.056 | Dominated |

The ERG's preferred base case was tested in sensitivity analyses for alemtuzumab versus SC IFNβ-1a 44µg (as this was the treatment used in the manufacturer's trials) for uncertainties that arose in the ERG critique of the manufacturer's model. All analyses shown below use the ERG's preferred base case.

**Table 37: Incremental deterministic base-case results of alemtuzumab versus SC IFNβ-1a 44 µg for the ERG's preferred base case**

| | SC IFNβ-1a 44 µg | | Alemtuzumab | | | | |
|---|---|---|---|---|---|---|---|
| **Technologies** | **Total costs (£)** | **Total QALYs** | **Total costs (£)** | **Total QALYs** | **Inc costs (£)** | **Inc QALYs** | **ICER (£) inc (QALYs)** |
| ERG preferred base case | 431,896 | 5.205 | 430,241 | 7.147 | -1655 | 1.942 | -852[a] |
| 50% reduction in disease progression natural history | 406,905 | 6.109 | 406,121 | 8.018 | -784 | 1.909 | -411[a] |
| Quality of life utility values LCI | 431,896 | 1.088 | 430,241 | 3.090 | -1655 | 2.002 | -827[a] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Quality of life utility values UCI | 431,896 | 9.038 | 430,241 | 10.975 | -1655 | 1.937 | -855[a] |
| Disease health state costs from Karampampa *et al.* 2012[45] | 548,917 | 5.205 | 506,465 | 7.147 | -42,452 | 1.942 | -21,862[a] |
| Disease health states costs from Biogen *et al*.2007 | 248,579 | 5.205 | 257,617 | 7.147 | 9038 | 1.942 | 4,654 |
| Relapse cost for hospitalisation, £3039 | 423,393 | 5.205 | 425,360 | 7.147 | 1966 | 1.942 | 1,013 |
| Waning effect, years 10+ 75% | 432,150 | 5.193 | 435,268 | 6.911 | 3117 | 1.718 | 1,815 |
| Waning effect, years 6-9 75%, 10+ 50% | 432,843 | 5.162 | 443,079 | 6.560 | 10236 | 1.399 | 7,319 |
| % patients receiving alumtuzumab, year 3 60%, year 5+ 20% | 431,896 | 5.205 | 446,160 | 6.917 | 14263 | 1.711 | 8,336 |
| MTC All years, 80% RRMS | 428,073 | 5.341 | 422,632 | 7.524 | -5440 | 2.184 | -2,491[a] |
| Disease progression using 6 month SAD | 437,211 | 4.936 | 426,446 | 7.333 | -10764 | 2.396 | -4,492[a] |

[a] Alemtuzumab dominates SC IFNβ-1a 44 µg, i.e. is cheaper and more effective.

There has been much criticism of the London Ontario dataset used for disease progression and that patients progress to more severe disease too rapidly.[26] The ERG considers that the London Ontario dataset to be unrepresentative of patients who may be given alemtuzumab or alternative treatments. We have varied the disease progression by reducing the transition probabilities to more severe health states by 50%. This analysis makes little difference to the cost effectiveness results (alemtuzumab continues to dominate).

In the previous appraisal for teriflunomide, the York ERG[35] explored the HRQoL estimates in the literature and considered there were wide variation around the estimates for the more severe health sates (i.e. 8 and 9). We ran the analyses using the 95% CI intervals for the Orme and colleagues[40] data used in the model. This showed that changing the quality of life values had minimal impact on the model results (alemtuzumab continue to dominate for both upper and lower confidence intervals).

There is a large variation in health costs for the source identified by Karampampa and colleagues[45] and Biogen and colleagues. We run sensitivity analyses using these health costs. There was wide variation in cost effectiveness results, with the ICER varying between £-21,862 for Karampampa and colleagues[45] (Alemtuzumab dominates) to £4,654 per QALY for Biogen.[50]

There was some uncertainty around the relapse cost for hospitalisation, with other alternative sources (Tyas and colleagues,[44] Kobelt and colleagues[51]). We used the cost recommended by the ERG for Dimethyl fumarate of £3039. This produced an ICER for alemtuzumab versus SC IFNβ-1a 44 µg of £1013 per QALY gained.

The GMS model assumes that the treatment effect continues at the same rate long term. However, it is plausible that the treatment effect wanes beyond the length of the clinical trial. We considered two scenarios with a waning effect. Where 100% treatment effect was assumed for years 1 - 10, and 75% treatment effect for years 10+, the ICER increased to £1815 per QALY for alemtuzumab versus SC IFNβ-1a 44 µg of £1013 per QALY gained. Where 100% treatment effect was assumed for years 1 - 5, 75% treatment effect for years 6 - 9, and 50% treatment effect was assumed for years 10+, the ICER for alemtuzumab versus SC IFNβ-1a 44 µg increased to £7,319 per QALY gained.

We investigated the sensitivity of the model results to changes in the proportion who receive subsequent doses of alemtuzumab. If the proportion of patients receiving alemtuzumab in year 3 were 60% and in years 5+ was 20%, the ICER increases to £8,336 per QALY gained.

The main source for the treatment effects used in the model is the manufacturer's post-2000 MTC and these effects were applied to the natural history of progression and relapse. We investigated the effect on the model results of using the all years MTC. In this case, the alemtuzumab continues to dominate SC IFNβ-1a 44 µg, with a more favourable ICER.

The 3-month SAD HRs results from the MTC were used in the manufacturer's model. However, it may be more appropriate to have used the 6-month SAD HRs. We have completed a sensitivity analysis with the 6-month SAD HR. For this analysis alemtuzumab continues to dominate SC IFNβ-1a 44 µg, with a more favourable ICER.

Superseded – see Erratum

*Treatment naive and treatment experienced subgroups*

The ERG considers that it is inappropriate to combine the trial evidence (Section 3.1). We have calculated the cost effectiveness for each of the trials using the ERG preferred base case (see above) and the relative risk for ARR and three month SAD for alemtuzumab versus SC IFNβ-1a 44µg. The parameter values for SC IFNβ-1a 44µg are as in the base case and the ARR and three month SAD parameter values for alemtuzumab are estimated using the relative risks from the trials.

The results are shown in Table 38. The treatment naive group, alemtuzumab dominated SC IFNβ-1a 44µg using the effectiveness from the CAMS 223 trial and the cost effectiveness was £6392 per QALY gained for alemtuzumab versus SC IFNβ-1a 44µg for the CARE MS I trial. Alemtuzumab dominated SC IFNβ-1a 44µg for a pooled analysis (ERG meta-analysis) of the two trials. For the treatment experienced group, using effectiveness data from CARE MS II, the cost effectiveness was £2854 per QALY gained for alemtuzumab versus SC IFNβ-1a 44µg.

**Table 38: ERG Subgroup analyses using for treatment naïve and experienced subgroups**

| | SC IFNβ-1a 44µg | | Alemtuzumab | | | | |
|---|---|---|---|---|---|---|---|
| **Technologies** | **Total costs (£)** | **Total QALYs** | **Total costs (£)** | **Total QALYs** | **Incr. costs (£)** | **Incr. QALYs** | **ICER (£) inc. (QALYs)** |
| Treatment naïve | | | | | | | |
| CAMS 223 | 431,896 | 5.205 | 402,960 | 8.459 | -28937 | 3.253 | -8,894 (Dominated) |
| CARE MS I | 431,896 | 5.205 | 441,006 | 6.631 | 9110 | 1.425 | 6,392 |
| Pooled (CAM 223, CARE MS I) | 431,896 | 5.205 | 423,531 | 7.463 | -8365 | 2.257 | -3,705 (Dominated) |
| Treatment experienced | | | | | | | |
| CARE MS II | 431,896 | 5.205 | 436,592 | 6.851 | 4695 | 1.645 | 2,854 |

*Subgroups for RES and HA populations*

The GMS performed separate MTC analyses and subgroup economic analyses for the HA despite interferon use and RES subgroups. The ERG has re-run these analyses using the ERG's preferred base case, for a slower progression from RRMS to SPMS and for different patient characteristics (for the RES and HA subgroups respectively). The results are shown in

Table 39. These changes make only minimal changes to the model results and alemtuzumab continues to dominate fingolimod and natalizumab in these subgroup analyses.

**Table 39: ERG Subgroup analyses for HA despite interferon use and RES**

| Technologies | Total costs (£) | Total QALYs | Total costs (£) | Total QALYs | Incr. costs (£) | Incr. QALYs | ICER (£) inc. (QALYs) |
|---|---|---|---|---|---|---|---|
| HA despite interferon use | Fingolimod[a] | | Alemtuzumab | | | | |
| GMS original analysis | 501,581 | 3.151 | 492,374 | 4.327 | -9207 | 1.176 | -7,828 (Dominated) |
| ERG's preferred base case | 450,390 | 5.643 | 420,230 | 7.600 | -30160 | 1.957 | -15,411 (Dominated) |
| RES | Natalizumab | | Alemtuzumab | | | | |
| GMS original analysis | 536,379 | 3.750 | 490,016 | 4.419 | -46,363 | 0.669 | -69,309 (Dominated) |
| ERG's preferred base case | 501,724 | 6.681 | 412,722 | 7.919 | -89,002 | 1.238 | -71,915 (Dominated) |

[a] Using an assumed PAS cost for fingolimod of £13,000

In summary, the ERG has tested the GMS model in a series of sensitivity analyses and has found the results robust to changes in assumptions and input parameters, with a cost effectiveness of less than £10,000 per QALY compared to SC IFNβ-1a 44 µg for all analyses.

### 4.4 Summary of uncertainties and issues

Clinical effectiveness issues

- The submission focuses on a more defined population, those with 'active' RRMS (participants had to have had a relapse within the last 12 months) than the NICE scope. Clinical advisors suggest this is appropriate.
- Trial populations did not fully meet their inclusion criteria for relapse.
- The population in one included trial had already failed to respond to a number of DMTs. The trial compared alemtuzumab to one of these DMTs and this should be taken into account when assessing the relative effectiveness of alemtuzumab compared with IFN-β 1a 44µg)
- There was the possibility of unmasking of outcome assessors across the three trials.
- There were different populations in the three pivotal trials, and the ERG does not consider it was appropriate to pool these data.

- Data were available to assess treatment naive and previously treated populations in subgroups as per the NICE scope, however, this was not undertaken.
- The only head to head comparison was with IFN-β1a. No head to head trials of alemtuzumab with the other comparators have been undertaken and so assessment of effectiveness is based on indirect comparison.
- The MTC was conducted reasonably, however, there was limited discussion of the limitations of the analysis, especially with respect to the subgroups analysed. In addition, the exclusion criteria applied meant that some direct comparisons were ineligible for the MTC.
- There is the possibility that some relevant studies were not included in the MTC because of the search cut-off used by the GMS.

Cost effectiveness issues

- Two subgroups not considered in the economic evaluation are treatment naïve and treatment experienced patients.
- There are limitations associated with both methods of estimating disease progression. Overall the ERG considers that the most appropriate method for the base case would be to use natural history from the placebo arms of trials such as the TEMSO and TOWER trials as these allow patients to improve their EDSS scores (although results are uncertain as they are for only a short time period), and the baseline characteristics from the CARE-MS trials,[3;4] because these patients would be more representative of those who would be offered treatment. However, these issues affect all treatments and the ERG additional analyses suggest that the impact on cost-effectiveness is probably limited.
- No clear explanation is given in the GMS or the choice of studies informing the HRQoL estimates for the economic evaluation, although both have been incorporated in previous STAs relating to MS. The ERG has some concerns about the use of these studies as the source of data seems arbitrary and CARE MS trial data could have been used, if data by health state were available. Also there are issues with one of the studies as mentioned in the previous assessments and the EQ-5D may not reflect changes well for MS patients. However, ERG analyses showed that changing the quality of life values had minimal impact on the model results.
- There is some uncertainty around the correct value for health state costs. As details of the constituents of the direct costs are not given, it is not clear what is included and

therefore it is difficult to assess if the approach is consistent with the NHS and PSS perspective. Sensitivity analyses run by the ERG showed a wide variation in cost-effectiveness results but alemtuzumab continued to dominate SC IFNβ-1a 44 µg.

# 5 End of life

NICE end of life treatment criteria were not applicable and not included in the MS.

# 6 Innovation

The GMS considers alemtuzumab as providing a 'step change' in patient care because it a) shows improved efficacy on disability accumulation endpoints against and active comparator, b) demonstrated reversal in mean EDSS scores against baseline in all studies. They also make the case that a high proportion of patients after the initial 2 year treatment course did not require retreatment and that this compares more favourably with other ongoing DMTs.

# 7 DISCUSSION

## 7.1 Summary of clinical effectiveness issues

The GMS includes evidence on the clinical effectiveness of alemtuzumab for active RRMS, including three RCTs of relevance to the scope. Results presented in the GMS suggest that on the co-primary outcome of SAD at 6-months alemtuzumab favoured treatment with IFN-β1a in two of the three trials. On the co-primary outcome of rate of relapse in the three included RCTs treatment with alemtuzumab statistically significantly reduced the rate of relapse when compared with IFNβ-1a. Key issues for consideration are the appropriateness of the pooling of the three RCTs, the comparator population in one trial receiving treatment they had already failed to respond to, limited data relating to the subgroups defined in the NICE scope, and some direct evidence missing from the MTC.

## 7.2 Summary of cost effectiveness issues

The GMS includes evidence on the cost-effectiveness of alemtuzumab for active RRMS compared to beta-interferons, glatiramer acetate, fingolimod and natalizumab. The model structure and methods adopted for the economic evaluation are reasonable and generally appropriate. The model structure is consistent with the clinical disease pathways and available

clinical trial evidence. The approach taken for model structure, assumptions and model parameter inputs follows that taken from previous economic models submitted to NICE technology appraisals for MS.

There are some areas of uncertainty relating to the long term modelling of MS. The population used in the model was based upon the UK RSS population, rather than the clinical trials' patient characteristics. In common with previous NICE MS technology appraisals, the GMS has derived natural history transition probabilities from the London Ontario dataset that are likely to overestimate the disease progression. There is also uncertainty around the correct value for health state costs.

# 8 REFERENCES


1. CAMMS223 Trial investigators, Coles AJ, Compston DA, Selmaj KW, Lake SL, Moran S *et al*. Alemtuzumab vs. interferon beta-1a in early multiple sclerosis. *NEJM* 2008;**359**:1786-801.

2. Cohen JA, Coles AJ, Arnold DL, Confavreux C, Fox EJ, Hartung HP *et al*. Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. *Lancet* 2012;**380**:1819-28.

3. Coles AJ, Twyman CL, Arnold DL, Cohen JA, Confavreux C, Fox EJ *et al*. Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. *Lancet* 2012;**380**:1829-39.

4. Coles AJ, Fox EJ, Vladic A, Gazda SK, Brinar V, Selmaj KW *et al*. Alemtuzumab more effective than interferon beta-1a at 5-year follow-up of CAMMS223 clinical trial. *Neurology* 2012;**78**:1069-78.

5. European Medicines Agency. *Draft guideline on clinical investigation of medicinal products for the treatment of Multiple Sclerosis*. European Medicines Agency; 2012. Available at: *http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/10/WC500133438.pdf*. Accessed

6. Fox EJ, Arnold DL, Cohen JA, et al. Durable Efficacy of Alemtuzumab in Relapsing-Remitting Multiple Sclerosis Patients Who Participated in the CARE-MS Studies: Three Year Follow-Up. *Neurology* 2013;**S41.001**.

7. Benedetti MD, Massacesi L, Tramacere I, Filippini G, La ML, Solari A *et al*. Non-inferiority of azathioprine versus interferon beta for relapsing remitting multiple sclerosis: A multicenter randomized trial. *Neuroepidemiology* 2012;**Conference**:3-4.

8. DeStefano N, Sormani MP, Stubinski B, Blevins G, Drulovic JS, Issard D *et al*. Efficacy and safety of subcutaneous interferon -1a in relapsing-remitting multiple sclerosis: further outcomes from the IMPROVE study. *Journal of the Neurological Sciences* 2012;**312**:97-101.

9. Hauser S, Kappos L, Li D, Calabresi P, O'Connor P, Bar-Or A *et al*. Long-term efficacy and safety of ocrelizumab in patients with relapsing-remitting multiple sclerosis (RRMS): Week 96 results of a phase II, randomized, multicenter trial. *Neurology* 2012;**Conference**:1.

10. Khan O, Rieckmann P, Boyko A, Selmaj K, Zivadinov R, GALA Study Group. Three times weekly glatiramer acetate in relapsing-remitting multiple sclerosis. *Annals of Neurology* 2013;**73**:705-13.

11. Massacesi L, Filippini G, Benedetti MD, La ML, Solari A, Tramacere I *et al*. Non-inferiority of azathioprine versus Beta interferons for relapsing remitting multiple sclerosis patients: A multicenter randomized trial. *European Journal of Neurology* 2012;**Conference**:28.

12. Nicholas R, Straube S, Schmidli H, et al. Trends in annualized relapse rates in relapsing-remitting multiple sclerosis and consequences for clinical trial design. *Mult.Scler.* 2011;**17**:1211-17.

13. Woods,.et al. Network meta analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. *BMC Medical Research Methodology* 2010;**10**:54.

14. Gupta, A., Arnold, D. L., Cohen, J., and et al.for the CARE-MS I investigators. Alemtuzumab Improves Quality Of Life Compared To SC IFNB-1A In CARE-MS I. Poster presented at CMSC, San Diego . 2012.

15. Arroyo, R., Arnold, D. L., Cohen, J. A., Confavreux, C., Fox, E. J., Hartung, H.-P., and et al.Compston for the CARE-MS II Investigators. Alemtuzumab Improves Quality Of Life Compared To SC IFNB-1a In CARE-MS II. Poster presented at the 23rd ENS, Spain . 2013.

16. Selmaj, K., Arnold DL, Cohen JA, and et al.Compston for the CARE-MS I and II Investigators. Alemtuzumab improves patient-reported quality of life in relapsing-remitting multiple sclerosis: CARE-MS I and II phase 3 trials. Poster presented at 22nd ENS, Prague P338. 2012.

17. Wingerchuk on behalf of the CAMMS223 Study Group. Efficacy of alemtuzumab in highly active relapsing-remitting multiple sclerosis patients in the CAMMS223 Trial. Presented at the 26th ECTRIMS & 15th RIMS,Sweden. Mult.Scler. (P49) S41-S196. 2010.

18. Krieger, S., Arnold D.L., Cohen, J., and et al.on behalf of CARE-MS II Investigators. Alemtuzumab Is Efficacious In Highly Active RRMS Patients In CARE-MS II. slides ref: at CMSC 2013, DX01 . 2013.

19. Confavreux, C., Twyman, C. L., and Arnold D.L. Efficacy of Alemtuzumab in Relapsing-Remitting Multiple Sclerosis (RRMS) Patients with Highly-Active Disease Despite Therapy. Presented at 16th EFNS, Sweden. European Journal of Neurology 19 (Suppl.1), 22-89. 2012.

20. Chilcott J, McCabe C, Tappenden P, O'Hagan A, Cooper NJ, Abrams K *et al*. Modelling the cost effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis. Commentary: evaluating disease modifying treatments in multiple sclerosis. *BMJ* 2003;**326**:522.

21. Association of British Neurologists. Association of British Neurologist's Guidelines for Prescribing in Multiple Sclerosis. 2009.

22. Gani R, Giovannoni G, Bates D, Kemball B, Hughes S, Kerrigan J. Cost-effectiveness analyses of natalizumab (Tysabri) compared with other disease-modifying therapies for people with highly active relapsing-remitting multiple sclerosis in the UK. *Pharmacoeconomics.* 2008;**26**:617-27.

23. Drummond MF,.Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;**313**:275-83.

24. Thompson JP, Abdolahi A, Noyes K. Modelling the cost effectiveness of disease-modifying treatments for multiple sclerosis: issues to consider. *Pharmacoeconomics.* 2013;**31**:455-69.

25. Yamamoto D,.Campbell JD. Cost-effectiveness of multiple sclerosis disease-modifying therapies: a systematic review of the literature. *Autoimmune.Dis.* 2012;**2012**:784364.

26. Hawton A, Shearer J, Goodwin E, Green C. Squinting through layers of fog: assessing the cost effectiveness of treatments for multiple sclerosis. *Appl.Health Econ.Health Policy* 2013;**11**:331-41.

27. NICE. *Multiple sclerosis (relapsing-remitting) - fingolimod: appraisal consultation document. TA254* . 2011. Available at: *http://guidance.nice.org.uk/TA/Wave20/71/ACD1/ACD1*. Accessed

28. NICE. *Natalizumab for the treatment of adults with highly active relapsing remitting multiple sclerosis. TA127* . 2007. Available at: *http://www.nice.org.uk/nicemedia/pdf/TA127Niceguidance.pdf*. Accessed

29. Weinshenker BG, Bass B, Rice GP, Noseworthy J, Carriere W, Baskerville J *et al*. The natural history of multiple sclerosis: a geographically based study. 2. Predictive value of the early clinical course. *Brain* 1989;**112 ( Pt 6)**:1419-28.

30. O'Connor P, Wolinsky JS, Confavreux C, Comi G, Kappos L, Olsson TP *et al*. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N.Engl.J.Med.* 2011;**365**:1293-303.

31. Office for National Statistics. England & Wales, Interim Life tables, 2008-2010. http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-223324 . 2011.

32. Pickin M, Cooper CL, Chater T, O'Hagan A, Abrams KR, Cooper NJ *et al*. The Multiple Sclerosis Risk Sharing Scheme Monitoring Study--early results and lessons for the future. *BMC.Neurol.* 2009;**9**:1.

33. Pokorski RJ. Long-term survival experience of patients with multiple sclerosis. *J.Insur.Med.* 1997;**29**:101-6.

34. Boggild M, Palace J, Barton P, Ben-Shlomo Y, Bregenzer T, Dobson C *et al*. Multiple sclerosis risk sharing scheme: two year results of clinical cohort study with historical comparator. *BMJ* 2009;**339**:b4677.

35. Evidence review Group. *ERG report: Teriflunomide for treating relapsing forms of multiple sclerosis*. NICE; 2013.

36. Katrych O, Simone TM, Azad S, Mousa SA. Disease-modifying agents in the treatment of multiple sclerosis: a review of long-term outcomes. *CNS.Neurol.Disord.Drug Targets.* 2009;**8**:512-9.

37. Held U, Heigenhauser L, Shang C, Kappos L, Polman C. Predictors of relapse rate in MS clinical trials. *Neurology* 2005;**65**:1769-73.

38. Patzold U,.Pocklington PR. Course of multiple sclerosis. First results of a prospective study carried out of 102 MS patients from 1976-1980. *Acta Neurol.Scand.* 1982;**65**:248-66.

39. Dee A, Hutchinson M, De La Harpe D. A budget impact analysis of natalizumab use in Ireland. *Ir.J.Med.Sci.* 2012;**181**:199-204.

40. Orme M, Kerrigan J, Tyas D, Russell N, Nixon R. The effect of disease, functional status, and relapses on the utility of people with multiple sclerosis in the UK. *Value Health* 2007;**10**:54-60.

41. Evidence review Group. *ERG report: natalizumab for the treatment of adults with highly active relapsing remitting multiple sclerosis.* NICE; 2007.

42. Naci H, Fleurence R, Birt J, Duhig A. The impact of increasing neurological disability of multiple sclerosis on health utilities: a systematic review of the literature. *J.Med.Econ.* 2010;**13**:78-89.

43. Evidence review Group. *ERG report: fingolimod for the treatment of relapsing remitting multiple sclerosis.* NICE; 2011.

44. Tyas D, Kerrigan J, Russell N, Nixon R. The distribution of the cost of multiple sclerosis in the UK: how do costs vary by illness severity? *Value Health* 2007;**10**:386-9.

45. Karampampa K, Gustavsson A, Miltenburger C, Tyas D. Treatment experience, burden and unmet needs (TRIBUNE) in MS study: results from the United Kingdom. *Mult.Scler.* 2012;**18**:41-5.

46. British National Formulary (BNF) (2013). Fingolimod/Gilenya information. Available at: http://www.medicinescomplete.com/mc/bnf/current/PHP5702-gilenya.htm#PHP5702-gilenya [Accessed 10 July 2013]. 2013.

47. British National Formulary (BNF) (2013). Natalizumab/Tysabri information. Available at: http://www.medicinescomplete.com/mc/bnf/current/PHP5725-tysabri.htm#PHP5725-tysabri [Accessed 10 July 2013] . 2013.

48. *Department of Health (2002). Health Service Circular HSC 2002/004, Cost Effective Provision Of Disease Modifying Therapies For People With Multiple Sclerosis - Risk Sharing Scheme for Disease Modifying Therapies in MS.* 2002.

49. National Institute for Health and Clinical Excellence. *Beta interferon and glatiramer acetate for the treament of multiple sclerosis. TA32*. 2002. Available at: *http://www.nice.org.uk/TA32*. Accessed

50. Biogen. *Natalizumab (Tysabri) for the treatment of adults with highly active relapsing remitting multiple sclerosis. Manufacturer submission (NICE Single Technology Appraisal submission)*. 2006.

51. Kobelt G, Berg J, Lindgren P, Kerrigan J, Russell N, Nixon R. Costs and quality of life of multiple sclerosis in the United Kingdom. *Eur.J.Health Econ.* 2006;**7 Suppl 2**:S96-104.