Evidence Review Group's Report Template

This template should be completed with reference to NICEs 'Guide to the Methods of Single Technology Appraisal'

Title: *Fingolimod for the treatment of relapsing remitting multiple sclerosis.*

Produced by *CRD and CHE Technology Assessment Group*

Authors *Miqdad Asaria , Research Fellow, CHE*
Gill Norman, Research Fellow, CRD
Sebastian Hinde Research Fellow, CHE
Joanne O'Connor, Research Fellow, CRD
Stephen Palmer, Professor of Health Economics, CHE
Kate Light, Information Specialist, CRD
Marta Soares, Research Fellow, CHE
Alison Eastwood, Senior Research Fellow, CRD

Correspondence to *Dr Gill Norman*
Centre for Reviews and Dissemination
University of York
York
YO10 5DD
09/06/2011

Date completed

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/247/01.

Declared competing interests of the authors
None

Acknowledgements
Dr Helen Ford, Department of Neurology, Regional Neurosciences Centre, *Leeds General Infirmary, Great George Street, Leeds, LS1 3EX*
Fiona Beyer, Research Fellow, CRD, University of York, York, YO10 5DD

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: *CRD and CHE Technology Assessment Group: Fingolimod for the treatment of relapsing remitting multiple sclerosis. Centre for Reviews and Dissemination/Centre for Health Economics; 2011*

Contributions of authors
Gill Norman wrote the clinical effectiveness sections of the report; Miqdad Asaria, Sebastian Hinde and Joanne O'Connor wrote the economic sections of the report;

Miqdad Asaria performed the economic modelling; Kate Light wrote the sections of the report dealing with search strategies and provided information support; Stephen Palmer commented on drafts of the economic sections; Marta Soares and Alison Eastwood managed the cost-effectiveness and clinical effectiveness sections of the project and reviewed and commented on the report.

Contents

List of Abbreviations

AE	adverse event
ALT	alanine aminotransferase
ARR	annualised relapse rate
BSC	best supportive care
CEAC	cost-effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
DMT	disease modifying therapy
EDSS	expanded disability status scale
EMA	European Medicines Agency
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D VAS	European Quality of Life-5 Dimensions visual analogue scale
ERG	evidence review group
GGT	gamma-glutamyl transferase
HR	hazard ratio
HRG	health resource group
HRQL	health related quality of life
ICER	incremental cost-effectiveness ratio
ITT	intention to treat
MRI	magnetic resonance imaging
MS	multiple sclerosis

MTC	mixed treatment comparison
NA	not applicable
NHS	National Health Service
PCT	Primary Care Trust
PRIMUS	Patient-Reported Indices for Multiple Sclerosis
PSA	probabilistic sensitivity analysis
PSS	personal social services
QALY	quality adjusted life year
RCT	randomised controlled trial
RES	rapidly evolving severe
RR	relative risk
RRMS	relapsing remitting multiple sclerosis
SPMS	secondary progressive multiple sclerosis
RSS	risk sharing scheme
UFIS	Unidimensional Fatigue Impact Scale

1 SUMMARY

1.1 Scope of the submission

This report presents the evidence review group (ERG)'s assessment of the manufacturer's (Novartis) submission to NICE on the use of fingolimod (Gilenya™) for the treatment of relapsing remitting multiple sclerosis (RRMS) in adults.

The final scope issued by NICE states that fingolimod is considered for adults with RRMS but that guidance will only be issued in accordance with the marketing authorisation. Committee for Medicinal Products for Human Use (CHMP) approval was subsequently issued for the following indications:

"As single disease modifying therapy in highly active RRMS for the following adult patient groups:

i) Patients with high disease activity despite treatment with a beta-interferon. These patients are defined as those who have failed to respond to a full and adequate course (normally at least one year of treatment) of beta-interferon. Patients should have had at least one relapse in the previous year while on therapy and have at least nine T2-hyperintense lesions in cranial magnetic resonance imaging (MRI) or at least one gadolinium-enhancing lesion. They may also be defined as patients with an unchanged or increased relapse rate or ongoing severe relapses compared to the previous year

ii) Patients with rapidly evolving severe (RES) RRMS defined by two or more disabling relapses in one year and with one or more gadolinium enhancing lesions on brain MRI or a significant increase in T2 lesion load as compared to a previous recent MRI."

1.2 Summary of submitted clinical effectiveness evidence

The manufacturer's submission focussed on the evidence from two phase three randomised controlled trials (RCTs) which directly compared fingolimod in its approved dose of 0.5mg/day with placebo (the FREEDOMS trial (N = 1,272))[1] and with interferon beta-1a (Avonex™) at 30 mcg/week (the TRANSFORMS trial (N =1,292)).[2] The primary outcome in both trials was annualised relapse rate (ARR). The duration of TRANSFORMS was 12 months; the duration of FREEDOMS was 24

months. Both trials also included a treatment arm of fingolimod 1.25 mg. The ARR for fingolimod 0.5 mg versus Avonex was 0.16 versus 0.33 ($p < 0.001$).[2] The ARR for fingolimod 0.5 mg versus placebo was 0.18 versus 0.40, ($p < 0.001$).[1] These relapse rates for comparator groups were low in both trials, relative to a general clinical population and to those of recent clinical trials. For example the AFFIRM trial had an ARR in the placebo group of 0.78 at 12 months and 0.73 at 24 months (compared to 0.27 and 0.23 respectively for the natalizumab group).[3]

The manufacturer identified (manufacturer's submission; table 47) the following post hoc subgroups within the FREEDOMS and TRANSFORMS trial populations as proxies for the patient groups (i) and (ii) (in section 1.1 above):

(i) Population 1 split into 1a and 1b:

Population 1a: Patients who were previously treated and have had at least one relapse in the prior year and *either* at least one gadolinium enhancing lesion *or* a T2 volume of greater than 0.5mL at baseline

Population 1b: Patients who were previously treated and have had equal of more relapses in year one than in year two

(ii) Population 2: Patients with two or more relapses and one or more gadolinium enhancing T1 lesions.

For populations 1a and 1b "previously treated" is defined as including treatment with glatiramer acetate as well as beta-interferon; this is justified on the basis that the EPAR states "intolerance to alternative MS therapy should also include Copaxone (glatiramer acetate) being tried". The use of 0.5mL T2 volume as a proxy for ≥9 T2 lesions is justified on pragmatic grounds of data availability and incorporation into the study database and the relationship between the criteria is explored.

The manufacturer's submission focussed on the approximation of population 1b identified above which was used as the base case for the submitted cost-effectiveness evidence. This represented 19.7% of the population of FREEDOMS and 43.6% of the population of TRANSFORMS. Outcome data were presented for the FREEDOMS and TRANSFORMS trial populations as a whole. Limited outcome data were also presented for population 1b for each trial. The ARR for fingolimod 0.5 mg versus Avonex in the approximation of population 1b was 0.25 versus 0.51 giving a ratio of 0.50 (95% CI 0.33 to 0.74) ($p < 0.001$), while that relative to placebo was

0.21 versus 0.54 giving a ratio of 0.38 (95% CI 0.24 to 0.62) ($p < 0.001$).[a] [redacted].

The ERG requested additional data on populations 1a, 1b and 2. There was significant overlap between these approximations of the CHMP patient groups. The ERG considered this to be potentially problematic, as fingolimod treatment for patients who meet criteria for population 2 would be most appropriately compared to treatment with natalizumab, rather than with Avonex. Data on populations 1a and 1b excluding patients who also met the criteria for population 2 were therefore also requested. After further discussion between the manufacturer and NICE the manufacturer subsequently provided, as a minimum, data on population 1b but not 2. There were no major differences in baseline characteristics between population 1b and population 1b but not 2. The ratio of ARR for population 1b but not 2 for fingolimod 0.5 mg versus Avonex up to month 12 was [redacted] while for fingolimod versus placebo up to month 24 it was [redacted].

In both trials serious adverse events were rare and were broadly comparable between the arms. While there were some differences in the incidence of specific adverse events, these generally followed a predicted pattern (for instance a lower incidence of influenza-type illness in fingolimod-treated patients compared to those in the Avonex arm of TRANSFORMS).

A mixed treatment comparison (MTC) was also included in the clinical effectiveness section of the submission. The MTC was comprised of 18 RCTs and included the following comparators: natalizumab, interferon beta-1a (Avonex, Rebif), interferon beta-1b (Betaferon (50 mcg/250 mcg)), glatiramer acetate (Copaxone), and placebo. The included trial populations were heterogeneous but broadly represented the RRMS patient population (with the exception of the trial comparing natalizumab with placebo[3]). The heterogeneity between trials was considered by the manufacturer to be substantial and the ERG agreed with this view. This remained the case even when consideration was limited to the trials which assessed fingolimod or Avonex. The MTC was not used to inform the economic model; instead an indirect

[a] Figures are as reported in the manufacturer's submission; the ERG was in some instances unable to reconcile the reported ARRs and the reported ratios of ARRs.

comparison using data from the FREEDOMS and TRANSFORMS trials was used to determine the efficacy of Avonex relative to placebo.

1.3 Summary of submitted cost-effectiveness evidence

The manufacturer undertook a systematic review of cost-effectiveness studies including fingolimod for the treatment of adult patients with RRMS. No cost-effectiveness evidence was found in the literature, thus a de novo economic model was developed. The model presented evaluates the cost-effectiveness of fingolimod compared with Avonex. No other comparators were considered. The original submission focused on population 1b. Results were not provided for any other sub-groups in the original submission.

The cost-effectiveness evaluation used a decision model, designed as a Markov model, to model disease progression using 21 health states representing different degrees of disease severity (by tracking EDSS scores whilst in RRMS and after conversion to SPMS) and death. Disability progression and conversion to SPMS were assumed irreversible. The model also accounted for relapses, adverse events, withdrawal and death. Although the occurrence of relapses did not influence the way in which progression was modelled to occur, relapse was modelled to depend on EDSS score. After withdrawing from fingolimod or Avonex, patients were assumed to receive best supportive care (BSC).

The perspective of the analysis of costs was that of the NHS and PSS. Costs were separated into disease costs, administration and monitoring costs and drug acquisition costs. QALYs were used as the measure of outcomes. Both patient and caregiver utility were accounted for and varied by disease severity. Utility adjustments were also applied to account for relapses and adverse events. Treatment with fingolimod or Avonex was assumed to be provided only to RRMS patients with an EDSS score of between 0.0 and 6.0. Patients were modelled to continue to receive these treatments until the treatment was either withdrawn (due to adverse events, disease progression to an EDSS score of above 6 or conversion to SPMS) or a patient died. A 50 year time horizon was used in the model to 'sufficiently capture differences in costs and outcomes;' previous appraisals in MS assessed by NICE adopted time-horizons lower than or equal to 20 years (TA 32 and TA 127).[4, 5] Both costs and benefits were discounted at 3.5%.

Natural history data were derived from external observational data sets and were used to inform the key events in the model. Natural history was assumed to

represent the course of disease under BSC; although the cost-effectiveness results of this treatment strategy were not presented in the submission, these were available in the model spreadsheet. External observational datasets were used despite data also being available from the clinical trials (FREEDOMS and TRANSFORMS) evaluating the use of fingolimod, Avonex and placebo. The treatment effects for fingolimod and Avonex for disability progression and relapse were derived from these two clinical trials; these were applied as relative risks to the appropriate summaries of the natural history data on the assumption that comparison with placebo in the FREEDOMS trial is representative of a comparison against BSC. Although FREEDOMS and TRANSFORMS followed patients for 24 and 12 months, respectively, in the model treatment effects were assumed to be sustained for as long as patients remained on treatment.

Data on mortality were derived from national mortality statistics and were adjusted for the additional risk of mortality for different EDSS states. Adjustment factors were derived from a combination of published studies (Pokorski (1997) and Sadovnik (1992)).[6, 7]

Both deterministic and probabilistic sensitivity analysis were carried out by the manufacturer to demonstrate the level of uncertainty around the model results. Despite the non-linear nature of the model, only the deterministic results were presented in the manufacturer's submission. The incremental cost-effectiveness ratio (ICER) of fingolimod relative to Avonex in population 1b was estimated to be £55,634 per QALY gained. The corresponding probabilistic estimate of the ICER (based on additional analysis carried out by the ERG maintaining the manufacturer's other assumptions for the base case) was estimated to be £69,787 per QALY gained.

In response to the ERG's request for further evidence, the manufacturer submitted additional cost-effectiveness analysis for population 1b but not 2 (based on the model described above). The model inputs modified to make predictions for this sub-population were patient characteristics and the estimates of the relative effectiveness of treatments. The ICER for fingolimod compared to Avonex was estimated to be £18,741 per QALY gained i.e. more favourable than in the previous analysis. The difference in the estimates was largely due to the revised relative efficacy estimates for Avonex suggesting that in this sub-population [redacted]. The relative risk of progression of Avonex vs. placebo was [redacted] in population 1b (i.e. approximately a [redacted]% [redacted] in the risk) and [redacted] (i.e. approximately a [redacted]% [redacted]) in population 1b but not 2. This additional analysis raised important

questions concerning the robustness of the model and the decision to restrict comparisons to be solely against Avonex.

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

Both the trials from which the head to head comparison data contained in the submission were drawn are good quality phase III RCTs which assessed fingolimod at its licensed dose in the RRMS population and were appropriately powered to assess the primary outcomes. Relevant primary outcomes were identified in the submission and consideration was given to secondary outcomes such as health related QoL. The comparators in these trials may be considered relevant: interferon beta-1a (Avonex) is one of the disease modifying therapies (DMT) used in treatment of RRMS under the NHS risk-sharing scheme, while placebo may be considered to be a proxy for BSC. The ERG's clinical advisor stated that patients with low EDSS scores enrolled in a placebo arm of a trial may receive a higher level of intervention than patients in the community treated with BSC; this would not be the case for patients with a higher EDSS score. The manufacturer has identified proxy populations for the indications for which fingolimod is licensed within the RRMS trial populations.

1.4.2 Weaknesses

The populations of both the trials which provide the direct evidence for the submission are adults with RRMS.[1, 2] The subpopulations which approximate the CHMP indications for use of fingolimod were defined post-hoc and do not precisely meet the CHMP criteria, although the ERG's clinical advisor has stated that these approximations are reasonable. More importantly, the extent of overlap between these subgroups is unclear in the original submission; clarification on this and, in particular, the extent to which population 2 (RES patients) are represented in populations 1a and, particularly, 1b was requested from the manufacturer. The initial submission provided baseline and outcome data only for population 1b; the ERG requested that this be supplied for populations 1a, 2, 1a but not 2, and 1b but not 2. The manufacturer subsequently provided data for population 1b but not 2.

Neither trial was conducted primarily in the UK. Both TRANSFORMS and FREEDOMS were multicentre and multinational trials. TRANSFORMS recruited patients from 172 centres in 18 countries in Europe, Asia, and North and South

America, while FREEDOMS recruited from 138 centres in 22 countries in Europe, the Middle East and southern Africa. Whilst the trial populations were over 90% Caucasian, they had lower mean EDSS scores (2.2 for TRANSFORMS[2] and 2.3 (fingolimod 0.5 mg)/2.45(placebo) for FREEDOMS[1]) than the population participating in the NHS risk sharing scheme (RSS)(mean EDSS for RRMS patients: 3.1[8]). However, the ERG's clinical advisor did not consider these differences likely to be clinically significant.

While the head-to-head comparisons address both an active comparator relevant to the NHS context and a proxy for BSC, there is no direct evidence presented for the comparison with either the other interferons used under the risk-sharing scheme, or for glatiramer acetate. Additionally, no direct evidence is presented for the comparison most relevant to population 2, which is with natalizumab;[9] patients who meet the criteria for population 2 were included in the base case population 1b as presented in the original submission. An MTC was presented in an attempt to provide data on efficacy and safety relative to these comparators but this had multiple limitations (see sections 4.1.7 and 4.2.2. below) and was not then subsequently used to inform the economic model.

The most serious concern with the choice of comparator is that beta-interferon treatment is used as a comparator in patients (population 1b) who, by definition have not responded to treatment with a DMT (beta-interferon). Therefore the Avonex arm represents continued use of a treatment which is ineffective in this group of patients; any estimate of relative efficacy derived from such a comparison is therefore highly likely to be non-conservative. This is supported by the fact that the ARR for the Avonex group in TRANSFORMS was 0.506 while that for the placebo arm of FREEDOMS was 0.542; the difference between these rates is very small indicating that the benefit over BSC conferred by Avonex may be extremely limited. Indeed the indirect comparison used in the economic model indicates that Avonex has [redacted] in population 1b but not 2, and is [redacted] than placebo. This is also indicative of the fact that it represents a non-ideal comparator.

Additionally, the active comparator for which a direct comparison is reported is Avonex, and the evidence for its efficacy relative to other beta-interferons is mixed, with conflicting results found for RCTs and non-randomised studies. Two systematic reviews comparing DMTs for RRMS were identified;[10, 11] only the review by Nikfar et al attempted a meta-analysis.[11] Whilst this meta-analysis of randomised and non-randomised studies showed no statistically significant difference in the pooled

relative risk (RR) of at least one relapse between Avonex and Rebif or between Avonex and interferon beta-1b (Betaferon), in each case the pooled estimate did not favour Avonex.[11] The RCTs included in the analysis favoured Rebif or Betaferon respectively. For a full consideration of the relative efficacy of the relevant comparators see section 3.3 below.

Evidence that Avonex may be less effective than other beta-interferons, suggests that a comparison solely against Avonex could lead to an overestimate of the cost-effectiveness of fingolimod. In the cost-effectiveness sections of the manufacturer's submission there is no attempt to assess the cost-effectiveness of fingolimod compared to other beta-interferons (or to other management options considered relevant comparators in the NICE scope) nor is there any attempt to consider the specific subgroup of patients with RES MS.

Overall the cost-effectiveness submission lacked clarity: descriptions of the methods used to generate input data are not explicit; calculations are not clearly explained and assumptions used are not justified.

The ERG has identified a number of key concerns in relation to the cost-effectiveness evidence submitted by the manufacturer. Comparisons with BSC are not presented in the submission. The ERG deems a comparison against BSC to be important since the sub-population considered in this analysis is one where patients have failed to respond to a previous course of DMTs. The cost-effectiveness of continued use of beta-interferon (or switching to an alternative product) in this subpopulation has not been evaluated in previous NICE appraisals and hence it should not be assumed that continued use of a beta-interferon is, in itself, cost-effective.

Other relevant comparators have also been excluded from the submission, despite being used in clinical practice. The use of Avonex as the main treatment comparator to fingolimod is not appropriately justified and the robustness of the cost-effectiveness results to the inclusion of alternative comparators is not considered.

There is limited justification for the data sources used to populate the key model parameters and inadequate description of the methods used and assumptions made in incorporating this data into the model.

The sensitivity analysis conducted is not adequately described in the submission and many key sources of uncertainty in the model have not been explored. The limited

analyses conducted show the results of the model to be highly sensitive to model parameters and assumptions. There is no justification of the choice of parameters used in the model in light of these uncertainties.

There is no attempt by the manufacturer to validate the model predictions either in terms of natural history or treatment effectiveness. The ERG's attempts to validate model predictions against trial data indicate that the model is unable to match the results observed in the trials.

The ERG's additional exploratory analysis has shown the sensitivity of the manufacturer's model to alternative sources of parameter data and alternative modelling assumptions. While the data sources selected and assumptions made have not been adequately justified by the manufacturer, the ERG has established that alternative choices of these lead to significant differences in the cost-effectiveness results estimated. In particular the ERG has shown that estimates of cost-effectiveness results are highly sensitive to changes in: the initial EDSS population distribution, interventions and comparators, natural history progression rates, waning of treatment effect, utility estimates, and the way effectiveness on relapse rates has been dealt with within the submission. This was observed for both the populations analysed, population 1b and population 1b but not 2.

1.4.3 Areas of uncertainty

There is considerable uncertainty as to the clinical and cost-effectiveness of fingolimod in the indicated populations relative to all comparators except beta interferon-1a (Avonex) and BSC for the primary outcome of ARR, even assuming that placebo represents a reasonable proxy for BSC. For the secondary outcomes of HR-QoL and MRI outcomes the efficacy of fingolimod relative to any comparator in the defined populations is also unclear. The efficacy of fingolimod relative to any comparator in populations 1a and 2 is also unclear.

It is unclear that Avonex is an appropriate comparator for fingolimod in population 1b, as these patients have, by definition, failed to respond to prior treatment with DMT which, in the great majority of cases, was beta-interferon. If interferon is accepted as an appropriate comparator for this base-case population then it is nonetheless unclear that an estimate derived from comparison with Avonex is conservative; the randomised evidence for the efficacy of Avonex relative to other formulations of interferon-beta suggests that this is not likely.

1.5 Key issues

The direct comparisons presented are with placebo which may be regarded as a proxy for BSC, and with interferon beta-1a (Avonex). Of the available DMT, it appears that Avonex is unlikely to be the most effective therapy. It also has relatively limited use in the NHS context; data supplied in the initial submission indicate use in 17.5% of patients approximating population 1b. The use of any beta-interferon as the main comparator for fingolimod in population 1b is also likely to be non-conservative, since population 1b have, by definition, failed to respond to prior DMT, which in almost all cases consisted of beta-interferon therapy.

An MTC was presented which attempts to assess efficacy and safety relative to other relevant comparators. Whilst there were high levels of heterogeneity between the included studies, which were in any case in the RRMS population as a whole, the fact that the results of this analysis were not then used to inform the economic model means that the uncertainty pertaining to the cost-effectiveness of fingolimod compared to relevant comparators remains. This is particularly the case given that the indirect comparison for population 1b but not 2 indicates that Avonex may be [redacted] than placebo, while Avonex is dominated or extendedly dominated in both populations 1b and 1b but not 2 (and is less cost-effective in population 1b than in population 1b but not 2).

Whilst the trials which form the primary focus of the submission were well conducted and adequately powered, they enrolled an RRMS population broader than the CHMP indicated populations. Approximations of these CHMP indicated populations were defined post-hoc, comprised a minority of trial participants, particularly in the case of the FREEDOMS trial, and included considerable overlap between the subgroups. Data for only one of these sub-populations were presented in the original submission (population 1b). As no data were presented for population 2, for whom the most relevant comparator is natalizumab, there is a large amount of uncertainty as to the relative efficacy of fingolimod and natalizumab in this population. Whilst the baseline characteristics of population 1b and 1b but not 2 did not differ substantially, the model outputs for the two populations did show substantial differences.

The MTC is based on trials which (with the exception of one trial comparing natalizumab versus placebo)[3] have RRMS populations, although there is considerable heterogeneity in the inclusion criteria. Therefore, any conclusions as to the relative efficacy of fingolimod compared to the comparators which are drawn from

the MTC must include considerable uncertainty as to their relevance to the CHMP indicated populations. As the MTC was not used to inform the economic model there is no assessment of the cost-effectiveness of fingolimod relative to any comparator other than Avonex.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer's description of the underlying health problem appeared appropriate and relevant and correctly characterised the different forms of the disease and the disease course which leads to the development of secondary progressive MS after a period of years with RRMS. The ERG's clinical advisor suggested that the reduced life expectancy of patients with MS is no longer so pronounced as was once the case, due to better multidisciplinary care in the later stages of the illness; it is not clear that this is reflected in the submission. Additionally, the ERG's clinical advisor stated that patients typically experienced more frequent relapses in the earlier stages of the disease course, and relapse frequency could change in an unpredictable manner during the disease course of a particular patient. She therefore advised us that the boundaries between groups which are partly defined by relapse frequency should be regarded as having a degree of fluidity; this was also not clear from the submission.

The number of patients estimated to be eligible for treatment with fingolimod was calculated by the manufacturer to be 6,697 patients based on the prevalence of MS in England and Wales, and the proportion of MS patients with RRMS who meet the criteria for their approximation of population 1b. This would represent 53% of the population of DMT-treated RRMS patients. The ERG's clinical advisor suggested that approximately 31% of the RRMS population would be eligible given that 10 to 15% of the total MS population (which includes patients who do not meet the criteria for RRMS) is currently treated with DMTs, with considerable regional variation. Therefore the manufacturer's estimates are not reasonable.

2.2 Critique of manufacturer's overview of current service provision

The overview of current service provision presented in the submission provided an accurate picture of the current availability of DMT both within and outwith the NHS RSS. The market shares (for England and Wales) of interferon beta-1a (Avonex and Rebif) and interferon-1b (Betaferon and Extavia) and glatiramer acetate were documented using prescribing data from the final quarter of 2010 for patients who would meet criteria for 1b. The manufacturer's response to the ERG's queries and clarifications presented data for patients in the general RRMS population which indicated a considerably higher market share for Avonex. The ERG takes the view that the data presented in the original submission gives a more accurate picture of the current service provision for the population(s) of interest, reflecting as it does prescriptions for patients who meet the criteria for the manufacturer's approximation of population 1b. It should be noted, however, that the initial submission did not take account of the fact that this approximation of population 1b included patients who would meet the criteria for population 2, and would hence meet the criteria for treatment with natalizumab.

While the description of current service provision was reasonable, it would have been helpful if there had been more consideration of the fact that the majority of patients are currently in receipt of BSC rather than DMT. The ERG's clinical advisor indicated that approximately 80% of patients do not currently receive any DMT. BSC will vary depending on the patient's individual needs. In patients with higher EDSS scores this may involve the participation of a multidisciplinary team which might include a specialist MS nurse, an occupational therapist and a physiotherapist as well as other professionals. BSC is also provided to patients in receipt of DMTs in addition to the clinic and/or hospital appointments directly related to their DMT. However, in the earlier stages of disease progression patients with low EDSS who are not taking DMT may require relatively limited support and/or health professional contact time, whereas those patients on DMT will have higher levels of contact as a consequence of the medication regimen.

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

3.1 Population

The NICE scope defined the population as adults with relapsing remitting multiple sclerosis (RRMS) but stated that guidance will only be issued in accordance with the marketing authorisation. This caveat acquired significance with the CHMP approval which was issued for highly active RRMS in the following adult patient groups:

1. Patients with high disease activity despite treatment with a beta-interferon. These patients are defined as:

   (a) those who have failed to respond to a full and adequate course (normally at least one year of treatment) of beta-interferon. Patients should have had at least one relapse in the previous year while on therapy and have at least nine T2-hyperintense lesions in cranial magnetic resonance imaging (MRI) or at least one gadolinium-enhancing lesion. They may also be defined as

   (b) patients with an unchanged or increased relapse rate or ongoing severe relapses compared to the previous year

2. Patients with rapidly evolving severe (RES) RRMS defined by two or more disabling relapses in one year and with one or more gadolinium enhancing lesions on brain MRI or a significant increase in T2 lesion load as compared to a previous recent MRI."[12]

The manufacturer defined post-hoc subgroups which approximated to the CHMP populations as follows.

Population 1a: Patients who were previously treated and have had at least one relapse in the prior year and *either* at least one gadolinium enhancing lesion *or* a T2 volume of greater than 0.5mL at baseline

Population 1b: Patients who were previously treated and have had equal or more relapses in year one than in year two

Population 2: Patients with two or more relapses and one or more gadolinium enhancing T1 lesions.

For populations 1a and 1b, “previously treated” is defined as including treatment with glatiramer acetate as well as beta-interferon; this is justified on the basis that the EPAR states “intolerance to alternative MS therapy should also include Copaxone (glatiramer acetate) being tried”. The use of 0.5mL T2 volume as a proxy for ≥9 T2 lesions was justified on pragmatic grounds of data availability and incorporation into the study database, and the relationship between the criteria was explored.

The ERG’s clinical advisor considered these to be reasonable approximations to the CHMP populations, being likely to capture the majority of eligible patients while not incorporating more than a minority of those ineligible under CHMP criteria. She did, however, note that since a criterion for treatment with DMT under the NHS RSS scheme was a minimum of two relapses in the previous two years, patients would have to meet a criterion of one relapse per year in the qualifying period. She also noted that it is difficult to comment on the relationship between the “severe relapses” defined in the CHMP population 1b and the “relapses” defined in the manufacturer’s approximation of this population.

The populations identified in the submission are not mutually exclusive. In particular a high proportion of patients meet the criteria for both population 1a and population 1b. Of greater concern is the fact that an unclear proportion of patients who meet the criteria for population 2 are included in populations 1a and 1b. Patients with RES RRMS are clearly a very different population, for whom Avonex is not the appropriate comparator; for this population the appropriate comparator is nataluzimab. Therefore the ERG requested that baseline and outcome data be supplied for populations 1a but not 2 and 1b but not 2 respectively. The initial response from the manufacturer did not provide these data; a subsequent request provided, as a minimum, data for population 1b but not 2.

Population 1b is already more likely than population 1a to show efficacy of fingolimod relative to interferon beta, as population 1a will include patients who have demonstrated some improvement on their previous DMT. The inclusion of patients meeting the criteria for population 2 in population 1b makes this population even less conservative and introduces uncertainty as to the true efficacy of fingolimod relative to appropriate comparators.

Comparison of baseline characteristics between population 1b and population 1b but not 2 revealed few differences between the populations; the fact that exploration of the model revealed significant differences in the cost effectiveness of fingolimod in the two populations suggests that it is highly sensitive to changes in parameters, and

that these parameters in turn are highly sensitive to minor changes in the patient population.

While there is no minimum EDSS score for the prescription of DMT's, patients in the NHS RSS scheme (who may be assumed to represent the UK RRMS population) have EDSS scores (mean: 3.1, SD 1.5 for RRMS patients) which are higher than the populations of both the FREEDOMS (mean 2.3, SD 1.3 for fingolimod/mean 2.5, SD 1.3 for placebo) and TRANSFORMS (mean 2.2, SD 1.3 for fingolimod/Avonex) trials. However, the ERG's clinical advisor did not consider this likely to have clinical significance, since the patients concerned had scores at the low end of the EDSS.

3.2 Intervention

Fingolimod has UK marketing authorisation for use in adults with RRMS who meet the criteria defined in section 3.1. The intervention described in the submission is that of oral fingolimod 0.5 mg/day which is in accordance with the NICE scope and with CHMP approval;[12] however the scope did not specify the dose at which fingolimod would be administered. As discussed in section 3.1 the submission defined the intervention in accordance with these indications. However, the FREEDOMS and TRANSFORMS trials employed it in populations with a broader diagnosis of RRMS.[1, 2] The submission identified post hoc subgroups within these trials which approximated to the populations indicated in the CHMP approval.

3.3 Comparators

The NICE scope defined the relevant comparators as being interferon beta, glatiramer acetate and optimised standard care with no DMT. For people with RES RRMS an additional comparator of natalizumab was identified. Interferon beta-1a and interferon beta-1b and glatiramer acetate are available under the NHS RSS which is operated in conjunction with the manufacturers of the relevant DMT. The exception to this is a form of interferon beta-1b (Extavia) which is not covered by the RSS; the Department of Health has advised that primary care trusts should be free to choose whether to use interferon beta-1b within (Betaferon) or outwith (Extavia) the RSS. Natalizumab is available to patients with RES RRMS as defined in the NICE guidance (2007).[9]

The manufacturer's submission identified the following comparators: Interferon beta-1a (Avonex, Rebif), interferon beta-1b (Betaferon, Extavia), glatiramer acetate

(Copaxone), BSC and, for patients with RES RRMS, natalizumab. Cladribine was excluded from the submission following the recent confirmation of the negative opinion by CHMP.

However, the trials which form the basis of the direct evidence in the submission assess only comparisons with placebo and interferon-1a (Avonex).[1, 2]

Use of Avonex within the NHS is relatively limited. The manufacturer's submission (Table A8) identifies it as being used in 17.5% of RRMS patients treated with DMT in England and Wales, compared with 36.1% treated with Rebif (44 mcg) and 25.8% treated with glatiramer acetate, based on prescribing data from the fourth quarter of 2010. Betaferon and lower dose (22 mcg) Rebif respectively accounted for 13.4% and 6.2% of patients treated and Extavia for only 1%. These figures represent data for patients who had previously been treated and had discontinued this treatment, and who had a stable or increased relapse rate over the previous year. Prescribing data for natalizumab which would be of relevance to population 2 were not presented. Given that over 80% of patients with characteristics approximating those of population 1b were treated with an alternative DMT, the fact that no evidence from head to head comparisons with these alternatives is available constitutes a clear weakness in the evidence base, which is compounded by the fact that there is mixed evidence as to the relative efficacy of Avonex compared to Rebif and Betaferon. While the submission is complete with respect to inclusion of extant head-to-head trials, it is reflective of this limited evidence base. A further issue arises from the fact population 1b, which formed the base case for the manufacturer's submission, also contains patients who meet the criteria for 2, and for whom natalizumab would therefore be the appropriate comparator intervention.

The manufacturer's response to the ERG's request for justification of the use of Avonex as a comparator included the provision of data obtained under the Freedom of Information Act which showed that Avonex was the most widely used DMT based on patients with RRMS in 60 primary care trusts from January 2008 to June 2010 (Response, p 17). Given that this is data for the general population of RRMS patients, the ERG's view is that the data presented in the original submission are more informative, as they approximate to population 1b rather than being based on the RRMS population as a whole. Furthermore, the ERG's clinical advisor considered that this subsequent data was incompatible with the overall prescribing data, whilst the PCTs from which the data were drawn were not reported.

The submission identified a number of trials which assessed head to head comparisons between the following comparators: interferon beta-1a 22mcg (Rebif); interferon beta-1a 44 mcg (Rebif); interferon beta-1a 30 mcg (Avonex); interferon beta-1b 250 mcg (Betaferon), glatiramer acetate (Copaxone) 20 mg; natalizumab 300 mg and placebo. These were combined in an MTC; however this was not subsequently used to inform the economic model, which therefore rests on the comparison with Avonex. Due to issues with the reported search strategy (see section 4.1.1 below), the ERG was unable to confirm that all relevant trials had been included. However, searches by the ERG did not reveal any additional RCTs which should have been included although a systematic review which included additional non-randomised studies was identified.[11] A second systematic review of RCTs which did not attempt statistical pooling was also identified, together with a Cochrane review which compared DMT to placebo.[10, 13]

The meta-analysis of randomised and non-randomised studies showed [redacted] in the pooled relative risk (RR) of at least one relapse between Avonex and Rebif or between Avonex and interferon beta-1b (Betaferon). However, in each case the pooled estimate [redacted] Avonex.[11] The RCTs included in the analysis [redacted]. The two RCTs included in the meta-analysis of Avonex versus Rebif[11] (the EVIDENCE trial and that of Etemadifar and colleagues (2006))[14,15] both showed statistically significant benefits of Rebif 44 mcg over Avonex for the outcome of experiencing at least one relapse, although the large cohort study of Limmroth and colleagues (2007) favoured Avonex.[16] Equally, while the four studies which were included in the meta-analysis of Avonex versus interferon beta-1b (Betaferon) for the outcome of experiencing at least one relapse showed contradictory results,[15-18] the single small RCT favoured Betaseron with an RR close to statistical significance (RR 0.72, 95% CI 0.48, 1.00).[15] Taken together, this evidence base provides support for the view of the ERG's clinical advisor, which was that Avonex was likely to be the least effective of the available formulations of beta-interferon.

However, the primary problem with the main comparator identified in the manufacturer's submission lies in the fact that it is one of the DMT to which the base-case population 1b is, by definition, resistant.

3.4 Outcomes

The primary outcomes identified in the manufacturer's submission are annualised relapse rate (ARR) and disability progression. Other outcomes reported are health-related quality of life (QoL) assessed using the EQ-5D and the EQ-5D VAS; MRI outcomes and adverse events including treatment discontinuation due to adverse events. The MTC was limited to ARR, confirmed disability progression at three months and discontinuation of treatment due to adverse events.

3.5 Time frame

The duration of both the trials assessing direct comparisons (TRANSFORMS and FREEDOMS) was short at 12 and 24 months respectively. In this respect they were comparable to other trials in the field; none of the trials included in the MTC had follow up exceeding 3.5 years. However, clearly the time frame for disease duration in MS is long term, with diagnosis typically between the ages of 20 and 40 years and life expectancy close to that of the general population. For patients with RRMS who are not treated with DMT, progression to SPMS typically occurs after an interval of between 5 and 20 years, with half of all patients progressing within 10 years of diagnosis with RRMS. The time-horizon for assessing impact on disease course is therefore very much longer than the available follow-up data from trial populations and in the model contained in the manufacturer's submission treatment effects are assumed to sustain for as long as patients are on treatment.

4 CLINICAL EFFECTIVENESS

4.1 Critique of manufacturer's approach

4.1.1 Description of manufacturers search strategy and comment on whether the search strategy was appropriate.

The submission was checked against the Single Technology Appraisal (STA) specification for manufacturer/ sponsor submission of evidence Update October 2009.

The manufacturer's submission described the search strategies used to identify relevant studies of fingolomid for the treatment of relapsing-remitting multiple sclerosis. Comparators searched for were: beta interferon, glatiramer acetate, natalizumab and standard care.

Search strategy for clinical evidence

The manufacturer's submission gave detailed descriptions of the search strategies and met NICE requirements. It included the specific databases searched (MEDLINE, MEDLINE In-Process, EMBASE and The Cochrane Library) and the service providers used, the date span of searches and the date searches were run. It also included the complete strategies used and the results for each set. The following Web sites were searched for conference abstracts that were published from 2008 to April 2010: American Academy of Neurology, Americas Committee for Treatment and Research in Multiple Sclerosis, European Committee for Treatment and Research in Multiple Sclerosis, European Charcot Foundation. Reference lists of the included studies and reviews were also searched for relevant studies.

There were some inappropriate elements in the search strategies used, such as the use of a facet to search the Cochrane Library for RCTs and use of economic studies search terms for NHS EED (these are inappropriate due to the content of the respective databases), and relevant material may have been missed as a consequence. However, the ERG did not identify any relevant studies which were not identified by the manufacturer's search. The search for clinical evidence may therefore be considered fit for purpose despite its non-ideal construction. As the searches for adverse events data, the MTC and non-RCT evidence employed the same strategy, they may also be considered fit for purpose, with the additional caveat that the use of a filter ir for both RCTs and non-RCTs (detailed in Tables 93 and 94 in the manufacturer's submission) may have contributed to relevant material being missed.

4.1.2 Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

The inclusion criteria used in the systematic review process were studies of patients with RRMS treated with the following interventions: fingolimod, any beta-interferon at all doses, glatiramer acetate, natalizumab and BSC. Mitoxantrone was excluded from the review. Trials of cladribine were included in the initial stage of study identification but subsequently excluded. RCTs, non-RCTs, long-term follow up studies and prospective observational studies, which were defined as phase IV studies were included. Relevant outcomes were relapse rate (mean ARR and patients remaining relapse free) disability progression (EDSS score and confirmed disability progression), disease activity, mortality, MRI measures, safety and tolerability

(including adverse event data and withdrawals from treatment) and health related quality of life. Immunology outcomes were excluded. Only studies reported in English were included in the review.

The dose of fingolimod was not specified in the inclusion criteria. However, the submission identified but subsequently excluded from consideration one RCT, and its extension studies, which assessed fingolimod at doses of 1.25 and 5.00 mg/day, above the licensed indication of 0.5 mg/day.[19] [20, 21] Whilst the relevant dose would ideally have been stated in the inclusion criteria, the criteria appeared appropriate to ensure the identification of relevant trials of fingolimod and appropriate comparators in the population defined in the NICE scope. The NICE scope is broader than the populations defined in the CHMP approval and the adoption of the wider criterion of RRMS was appropriate to capture all relevant studies.

4.1.3 Table of identified studies. What studies were included in the submission and what were excluded.

Evidence of direct comparisons between fingolimod and placebo and between fingolimod and Avonex came from the FREEDOMS and TRANSFORMS trials respectively.[1, 2] Details of the populations in these trials are shown in table 1 below. An additional trial was identified but was excluded as fingolimod was not assessed at the licensed dose of 0.5 mg, but only at the higher doses of 1.25 mg and 5.0 mg.[19] The ERG accepts that this decision was reasonable.

The trials included in the MTC are shown in table 2. This clearly shows that there were considerable differences in whether prior DMT was permitted, trial duration, and definition of disability progression, even if the analysis were restricted to trials which assessed Avonex and/or Fingolimod.

Table 1: Population characteristics of the FREEDOMS and TRANSFORMS trials.

	FREEDOMS			TRANSFORMS		
Trial Arm	Fingolimod 1.25 mg	Fingolimod 0.5 mg	Placebo	Fingolimod 1.25 mg	Fingolimod 0.5 mg	Interferon Beta 1a
N (ITT)	429	425	418	426	431	435
N (modified ITT)	NA	NA	NA	420	429	431
Age: years **Mean ± SD** **Median (range)**	 37.4 ± 8.9 38.0 (17-55)	 36.6 ± 8.8 36.0 (18-55)	 37.2 ± 8.6 37.0 (18-55)	 35.8±8.4 36 (18-54)	 36.7±8.8 37 (18-55)	 36.0±8.3 36 (18-55)
Females: N (%)	295 (68.8)	296 (69.6)	298 (71.3)	293 (68.8)	282 (65.4)	295 (67.8)
EDSS *Mean ± SD* *Median (range)*	 2.4 ±1.4 2.0 (0-5.5)	 2.3 ±1.3 2.0 (0-5.5)	 2.5 ± 1.3 2.0 (0-5.5)	 2.21 ± 1.2 2.0 (0-5.5)	 2.24 ±1.3 2.0 (0-5.5)	 2.19±1.26 2.0 (0-5.5)
Time from first MS symptom (yr) *Mean ± SD* *Median (range)*	 8.4 ± 6.9 6.9 (0-37)	 8.0 ± 6.6 6.6 (0-35)	 8.1 ± 6.4 7.0 (0-32)	 7.3± 6.0 6 (0-33)	 7.5±6.2 6 (0-34)	 7.4 ±6.3 6 (0-40)
History of DMT: None Any interferon beta Glatiramer acetate Natalizumab	259 (60.4)	244 (57.4)	249 (59.6)	177 (41.5) 209 (49.1) 67 (15.7) 3 (0.7)	193 (44.8) 219 (50.8) 57 (13.2) 4 (0.9)	190 (43,7) 207 (47.6) 67 (15.4) 1 (0.2)
Relapse history: N Within previous year *Mean ± SD* *Median (range)* Within previous 2 years *Mean ± SD* *Median (range)*	 1.5 ± 0.8 1.0 (0-6) 2.1 ± 1.3 2.0 (1-10)	 1.5 ± 0.8 1.0 (0-5) 2.1 ± 1.1 2.0 (1-11)	 1.4 ± 0.7 1.0 (0-6) 2.2 ± 1.2 2.0 (1-10)	 1.5 ±0.9 1 (0-7) 2.2±1.2 2 (1.8)	 1.5±1.2 1 (0-20) 2.3±2.2 2 (1-40)	 1.5±0.8 1 (0-6) 2.3±1.2 2 (1-12)

MRI	N = 424	N = 424	N = 416	412	427	425
Absence of gadolinium enhancing lesions: N (%)	257 (60.6)	263 (62.0)	262 (63.0)	270 (65.5)	288 (67.4)	268 (63.1)
No gadolinium enhancing lesions: N (%)						
Mean ± SD	1,8 ± 4,7	1.6 ± 5.6	1.3 ± 2.9	1.49±4.77	0.98± 2.81	1.06±2.80
Median (range)	0 (0-50)	0 (0-84)	0 (0-26)	(0-66)	0 (0-29)	0 (0-36)
Vol. hypointense lesions on T_2 weighted images (mm^3)						
Mean ± SD	6829 ± 8491	6128 ± 7623	1962±3131	5085±5962	5170±6642	4924±5711
Median (range)	3557 (0-47,734)	3303 (0-47,148)	811 (0-20,956)	3096 (0-38,870)	2382 (0-46,280)	2901 0-38,712)
Normalised brain vol. (ml)						
Mean ± SD	1511 ± 86	1521 ± 85	1512 ± 85	1526.2±76.4	1524.1±83.9	1526.7±77.9
Median (range)	1515 (1217-1764)	1529 (1144-1734)	1515 (1230 -1723)	1528(1300-1794)	1526 (1185-1862)	1533 (1231-1762)

Table 2: characteristics of trials included in the MTC

	Interventions	N	Duration	Disability progression	Mean age (yrs)	% female	Disease duration (yrs)	Relapses in last 2 yr	Mean EDSS	Prior DMT Permitted?
AFFIRM[3]	Natializumab vs placebo	942		EDSS+1 or 1.5 from baseline of 0 over 12 weeks	35.6 vs 36.7	71.6 vs 67.0	5.0 vs 6.0	3.31 vs 2.27 (imputed from 1 year data)	2.30 vs 2.30	None in previous 6 months, not for more than a total of 6 months
EVIDENCE[14]	Avonex vs Rebif	677	48 wks	EDSS+1 over 3 months	37.4 vs 38.3	74.6 vs 74.9	6.7 vs 6.5	2.27 vs 2.60	2.60 vs 2.60	No interferon
FREEDOMS [1]	Fingolimod vs placebo	1272	2 yr	EDSS+1 (or 0.5 if over 5.5) over 3 months	36.6 vs 37.2	69.6 vs 71.3	8.0 vs 8.1	2.10 vs 2.20	2.30 vs 2.50	Yes 62.6 vs 60.4
INCOMIN [22]	Avonex vs Betaferon	188	2 yr	EDSS+1 over 6 months	34.9 vs 38.8	62.0 vs 68.8	6.7 vs 5.9	2.76 vs 3.04	1.96 vs 1.97	No interferon or immunosuppressant except corticosteroids
MSCRG [23]	Avonex vs placebo	301	2 yr	EDSS+1 over 6 months	36.7 vs 36.9	74.7 vs 72.0	6.6 vs 6.4	1.81 vs 1.81 (imputed from 1 year data)	2.40 vs 2.30	No interferon or immunosuppressant
IFNB MS Study Group[24]	Betaferon 250 mcg vs placebo vs Betaferon 50 mcg	372	2 yr	EDSS+1 over 3 months	35.2 vs 36.0 vs 35.3 (median)	69.3 vs 71.5 vs 68.0	4.7 vs 3.9 vs 4.7	3.4 vs 3.6 vs 3.3	3.0 vs 2.8 vs 2.9	No aziothioprine or cyclophosphamide
PRISMS[25]	Rebif 22mcg vs placebo vs Rebif 44 mcg	560	2 yr	EDSS+1 over 90 days	34.8 vs 34.6 vs 35.6 (median)	67.2 vs 74.9 vs 65.8	5.4 vs 4.3 vs 6.4	3.0 vs 3.0 vs 3.0	2.5 vs 2.4 vs 2.5	No interferon, no other immunosuppressive treatment in prior 12 months
TRANSFORMS [2]	Fingolimod vs Avonex	1292	1 yr	EDSS+1 over 3 months	36.7 vs 36.0	65.4 vs 67.8	7.5 vs 7.4	2.3 vs 2.3	2.24 vs 2.19	Yes 55.2 vs 56.3
BEYOND[26]	Betaferon 250 mcg vs glatiramer acetate vs betaferon 500 mcg	2244	2-3.5 yr	EDSS+1 (or 0.5 if over 5.5) over 3 months	35.8 vs 35.2 vs 35.9	69.9 vs 68.3 vs 70.0	5.3 vs 5.1 vs 5.4	2.42 vs 2.42 vs 2.42 (imputed from 1 year data)	2.35 vs 2.28 vs 2.33	No
BECOME[27]	Betaferon 250 mcg vs glatiramer acetate	75	2 yr	EDSS+1 over 3 months	36.0 vs 36.0	75.0 vs 64.1	0.9 vs 1.2	2.72 vs 2.87 (imputed from 1 year data)	2.00 vs 2.00	NR
REGARD[28]	Rebif vs glatiramer acetate	764	96 wks	NA	36.7 vs 36.8	69.0 vs 72.0	5.93 vs 6.55	NR	2.35 vs 2.55	No interferon, glatiramer acetate or cladribine
Hurwitz 2008[29]	Betaferon	71	12-28	EDSS+1 (or 1.5	37.9 vs 37.8	71.0 vs 76.0	NR	NR	2.8 vs 2.0	NR

	250 mcg vs Betaferon 500 mcg		weeks	if 0, or 0.5 if ≥5.0) over 6 months					(median)	
Etemadifar 2006[15]	Avonex vs Rebif vs Betaferon	90	2 yr	NA	31.0 vs 30.4 vs 33.6 Calculated from age of onset + MS duration (submission mistakenly gives age of onset)	80 vs 76.7 vs 73.8	2.9 vs 3.0 vs 3.7	3.02 vs 3.02 vs 3.02 imputed from 1 year data	1.9 vs 2.1 vs 1.9	Yes Numbers not reported
Wroe 2005[30]	Betaferon 250 mcg vs placebo	98	90 days	NA	35.0 vs 38.0	73.8 vs 72.7	NR	2.66 vs 2.47	2.92 vs 3.09	No
Saida 2005[31]	Betaferon 250 mcg vs Betaferon 50 mcg	205	2 yrs	NA	35.5 vs 36.3	71.9 vs 66.7	6.30 vs 8.00	3.02 vs 2.87 (imputed from 1 year data)	NR	NR
Johnson 1995[32]	Glatiramer acetate vs placebo	251	2 yrs	NA	34.58 vs 34.33	NR	7.25 vs 6.64	2.91 vs 2.93	2.82 vs 2.42	NR
Comi 2001[33]	Glatiramer acetate vs placebo	239	9 mths	NA	34.1 vs 34.0	NR	7.90 vs 8.30	2.80 vs 2.50	2.30 vs 2.40	No
Bornstein 1987[34]	Glatiramer acetate vs placebo	48	2 yrs	1 unit on the Krutzke score over 3 months	30.0 vs 31.0	56.0 vs 60.0	4.90 vs 6.10	3.80 vs 3.90	2.90 vs 3.20	NR

This table was partially based on figures drawn from the main trial publications due to inconsistencies between trial identifiers and other errors in the manufacturer's submission.

4.1.4 Details of any relevant studies that were not included in the submission

The ERG did not identify any relevant completed studies which were not included in the submission. Although the Cochrane review included some additional placebo-controlled studies, these reported only secondary outcomes (e.g. MRI data).[13] Searches of clinical trials databases by the ERG revealed no relevant ongoing studies. As noted above, the decision to exclude from consideration a trial of fingolimod at doses above that considered in the CHMP was considered by the ERG to be appropriate.[19] The decision to exclude trials assessing cladribine from the MTC, following the negative opinion by CHMP was also considered by the ERG to be appropriate.

4.1.5 Description and critique of manufacturers approach to validity assessment

The manufacturer assessed the trials which evaluated direct comparisons with fingolimod against the following criteria:

The ERG's validity assessment of these studies is shown in table 3 below.

Table 3:

Trial number (acronym)	Study D2302 (TRANSFORMS)	Study D2301 (FREEDOMS)
Was randomisation carried out appropriately?	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Some evidence of greater burden of gadolinium enhancing lesions in fingolimod1.25 group compared to placebo or fingolimod 0.5
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes. There was a serious attempt to prevent adverse event profiles from revealing allocation to assessors for efficacy outcomes	Yes
Were there any unexpected imbalances in drop-outs between groups?	Slightly higher proportion of fingolimod 1.25mg patients dropped out due to adverse events	Lower proportion of fingolimod 0.5 patients discontinued than in placebo of fingolimod 1.25.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No

Trial number (acronym)	Study D2302 (TRANSFORMS)	Study D2301 (FREEDOMS)
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	No (modified ITT analysis used, including patients who were randomized and received at least one dose of fingolimod).	Yes
Was there an appropriate sample size calculation?*	Yes	Yes

*Not part of the current NICE criteria but relevant criterion.

The studies included in the MTC were appraised (manufacturer's submission, table 98) using the criteria of randomisation methods and blinding. Whilst these deal with some basic aspects of trial validity they do not consider allocation concealment, use of an intention-to-treat analysis, selective outcome reporting or the use of a power calculation. The results of the appraisal indicated that most trials were of reasonable quality, based on the criteria assessed.

The ERG did not replicate the validity assessment of the trials in the MTC; since the MTC is not used to inform the economic model this was not considered necessary.

4.1.6 Description and critique of manufacturers outcome selection

The primary outcomes identified in the manufacturer's submission are annualised relapse rate (ARR) and disability progression. Other outcomes reported are health-related quality of life (QoL) assessed using the EQ-5D and the EQ-5D VAS; MRI outcomes and adverse events including treatment discontinuation due to adverse events.

There are variations between trials in the way in which relapses, and hence ARR are defined. This is particularly the case with the trials included in the MTC. In particular the criteria for severe and disabling relapses as opposed to clinical relapses are unclear and in some instances are not defined.

As with the ARR there are differences between trials in the way in which disability progression is defined in relation to EDSS; these particularly relate to the time required to confirm progression as well as to the way in which relapses are defined by the trial (confirmation usually requires that the patient be relapse free at assessment). Again, these differences are particularly notable in the trials included in the MTC.

Adverse event data covered the range of outcomes reported in the trials and was sufficiently comprehensive to permit an appraisal of the safety profile of fingolimod 0.5 mg compared to both Avonex and placebo.

The use only of the EQ-5D and EQ-5D-VAS to assess HR-QoL in the FREEDOMS trial is unlikely to be ideal. Disease specific measures of QoL in MS patients exist and would have provided more relevant and detailed information on this outcome.[35] The use of the PRIMUS –QoL, PRIMUS – Activities and UFIS in the TRANSFORMS trial was perhaps more clinically informative but was not used to inform the economic model.

Data on a number of MRI measures of disease activity were presented, including the key outcomes of Gadolinium-enhancing lesions and T2-lesions.

The MTC was limited to ARR, confirmed disability progression at three months and discontinuation of treatment due to adverse events. The ERG's clinical advisor stated that discontinuation due to adverse events was a reasonable means of assessing acute adverse events but would not capture longer term adverse events, the assessment of which would be dependent on post-marketing surveillance.

4.1.7 Describe and critique the statistical approach used

The two trials of head-to head comparisons (TRANSFORMS and FREEDOMS) appraised fingolimod 0.5 mg and 1.25 mg compared, respectively to Avonex 30 mcg and to placebo. Consequently there was no attempt statistically to combine the efficacy data from these trials in the clinical effectiveness sections of the submission; this was clearly appropriate.

The fingolimod 0.5 mg arms were combined in the analysis of safety outcomes for comparison with data from the Avonex arm of the TRANSFORMS trial and the placebo arm of the FREEDOMS trial respectively. This appeared reasonable despite the differences between the trials.

The populations of both TRANSFORMS and FREEDOMS were broader than those for which CHMP indicated approval. Consequently the submission included post-hoc subgroups which approximated these CHMP populations. Post-hoc identification of subgroups has attendant problems which must be considered when evaluating the strength of the evidence represented by trials. These subgroups showed considerable degrees of overlap and the population which approximated 1b contained a significant number of patients who met the criteria for 2 as well as a majority of patients who were also included in population 1a. The subgroups were also relatively small; population 1b comprised fewer than half the population of TRANSFORMS while under 20% of the FREEDOMS population met the criteria for

population 1b. Whilst the trials were adequately powered to assess comparative efficacy in the whole RRMS populations recruited, there must be serious concern about their power to assess this in these relatively small post-hoc subgroups, particularly in the case of FREEDOMS.

Whilst the head-to-head comparisons reported HRs for efficacy outcomes, the MTC employed RR's. This is potentially problematic as, unlike an odds ratio, the RR is not symmetric. This fact has been demonstrated to be capable of generating anomalous results in an indirect comparison, including for an analysis comparing natalizumab with interferon therapy for the outcome of progression.[36] A further point to note is that the MTC was conducted using PROC GLIMMIX in SAS; this is known to incorporate lower levels of uncertainty around the means than WINBUGS. Therefore it is possible that the analysis may not reflect the full extent of the heterogeneity which was apparent between the trials.

4.1.8 Summary statement

The submission appears complete in its inclusion of extant RCTs assessing direct comparisons between fingolimod and relevant comparators. These trials assessed fingolimod in the RRMS population which is broader than the NICE scope following the CHMP opinion.[1, 2] The manufacturer subsequently defined post-hoc approximations of the populations for which the CHMP approval was issued within both trials. Whilst the ERG's clinical advisor considered (with some caveats) these approximations to be reasonable proxies for the CHMP populations, incomplete data were provided for the subgroups.

For population 1b which represents the manufacturer's base case data were provided for primary but not secondary outcomes. Data for populations 1a and 2 were not presented. The data for population 1b presented in the original submission included a proportion of patients who met the criteria for population 2, for whom an appropriate comparator would be natalizumab rather than Avonex or alternative interferons.

The ERG did not identify any relevant ongoing studies.

It is not clear whether all relevant studies were identified in the MTC (see section 4.1.1 above); the ERG did not have the resources to replicate the searches, but nor is it aware of any additional relevant studies which should have been included. Since the MTC is not used to inform the economic model this was considered of secondary importance.

The lack of direct evidence on efficacy and safety relative to relevant comparators other than Avonex and placebo (see discussion elsewhere on relationship between placebo and BSC) is a clear weakness in the evidence base on which the submission rests. A consequence of

this lack of head-to-head comparisons is the high level of uncertainty which pertains to the efficacy of fingolimod 0.5 mg relative to any other comparator. The MTC which the manufacturer's submission presented primarily rested on heterogeneous trials in the general RRMS population and as a consequence was not used to inform the economic model. These weaknesses in the evidence base are accentuated by the fact that the balance of evidence suggests that Avonex may not represent the most effective available interferon therapy. The final, and most serious problem lies with the fact that patients in population 1b have, by definition, failed to respond to DMT (interferons in the vast majority of cases). The use of any interferon as a comparator in this population therefore may be considered to be non-ideal. This view is supported by the similarity of the ARR in the Avonex arm of TRANSFORMS to that of the placebo group in FREEDOMS (see section 4.2.1 below). Any estimate of relative efficacy derived from such a comparison is therefore likely to be non-conservative.

4.2 Summary of submitted evidence

4.2.1 Summary of results

The initial submission contained results for the whole trial populations and, in the case of primary outcomes, results for population 1b. The company subsequently supplied outcome data for population 1b but not 2.

Annualised Relapse Rate

In the TRANSFORMS trial the ARR up to month 12 for fingolimod 0.5 mg versus Avonex 30 mcg was 0.16 versus 0.33 ($p < 0.001$).[2] The ARR for fingolimod 0.5 mg versus Avonex in the approximation of population 1b was 0.25 versus 0.51 giving a ratio of 0.50 (95% CI 0.33 to 0.74 ($p < 0.001$).[b] In population 1b but not 2 the ARR was 0.25 (95% CI 0.17 to 0.35) for fingolimod 0.5 mg versus 0.44 (95% CI 0.33 to 0.59) for Avonex 30 mcg, with a ratio of ARR of 0.44 (95% CI 0.31 to 0.64).

In FREEDOMS the ARR up to month 24 for fingolimod 0.5 mg versus placebo was 0.18 versus 0.40, ($p < 0.001$).[1] The ARR for fingolimod 0.5 mg versus placebo in the approximation of population 1b was 0.21 versus 0.54 giving a ratio of 0.38 (95% CI 0.24 to 0.62) ($p < 0.001$). In population 1b but not 2 the ARR was 0.19 (95% CI 0.12 to 0.29) for

[b]Figures are as reported in the manufacturer's submission; the ERG was in some instances unable to reconcile the reported ARRs and the reported ratios of ARRs.

fingolimod 0.5 mg versus 0.44 (95% CI 0.31 to 0.63) for placebo, with a ratio of ARR of 0.45 (95% CI 0.35 to 0.57). The differences between populations 1b and 1b but not 2 are summarised in table 4

Table 4: ARR for population 1b and population 1b but not 2 in the TRANSFORMS and FREEDOMS trials

	TRANSFORMS (12 months)			FREEDOMS (24 months)		
	ARR: fingolimod 0.5 mg (95% CI)	ARR: Avonex 30 mcg (95% CI)	Ratio of ARR (95% CI)	ARR: fingolimod 0.5 mg (95% CI)	ARR: placebo (95% CI)	Ratio of ARR (95% CI)
Population 1b	0.25 (CI not reported)	0.51 (CI not reported)	0.50 (0.33 to 0.74)	0.21(CI not reported)	0.54 (CI not reported)	0.38 (0.24 to 0.62)
Population 1b but not 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Disability Progression

In the TRANSFORMS trial the proportion of patients with no disability progression at 12 months was 94.1% (95% CI: 91.8 to 96.3) for fingolimod 0.5 mg versus 92.1% (95% CI 89.4 to 94.7) for Avonex 30 mcg. The HR for disability progression for population 1b was reported as [redacted]. For the population of 1b but not 2 the HR was [redacted].

In the FREEDOMS trial the proportion of patients with no disability progression at 24 months was 82.3% (95% CI 78.6 to 86.1) for fingolimod 0.5mg versus 75.9% (95% CI 71.7 to 80.2) for placebo, giving an HR of 0.70 (0.52 to 0.96). In population 1b the HR was [redacted]. For the population of 1b but not 2 the HR was [redacted].

Health-related quality of life

The TRANSFORMS study assessed patient-reported outcomes using the Patient-Reported Indices for Multiple Sclerosis – Quality of life (PRIMUS - QoL); the Patient-Reported Indices for Multiple Sclerosis – Activities (PRIMUS – Activities) and the Unidimensional Fatigue Impact Scale (UFIS). There were no statistically significant differences between the Fingolimod 0.5 and Avonex groups in change from baseline on the PRIMUS-QoL or the UFIS. The PRIMUS – Activities scale showed a statistically significant benefit of fingolimod on changes in ability to perform daily activities (fingolimod 0.08 ± 4.47 versus Avonex 0.43 ± 4.71; $p < 0.05$).

In the FREEDOMS study patient-reported outcomes were assessed using the EQ-5D, and [redacted]

MRI outcomes

In the TRANSFORMS trial, the fingolimod 0.5 mg group had significantly less disease activity than the Avonex group activity as assessed by a number of parameters. These included the number of new or enlarged hyperintense lesions on T2-weighted images and number of gadolinium-enhancing lesions on T1-weighted images (there were no significant differences in volume of gadolinium-enhancing lesions). Statistically significantly more patients in the fingolimod 0.5 mg group were free from MRI activity compared to those in the Avonex group. There was also a significantly lower reduction from baseline in brain volume in the fingolimod group (see manufacturer's submission, table 26). The FREEDOMS trial also found benefits of fingolimod 0.5 mg over placebo on a range of MRI measures of disease activity (see manufacturer's submission, table 28).

Adverse events

The majority of adverse events assessed in the FREEDOMS and TRANSFORMS trials showed no statistically significant differences between the Fingolimod 0.5 mg and placebo or Avonex 30 mcg arms respectively. The submission combined the fingolimod 0.5 mg arms from the two trials for the assessment of safety outcomes. The ERG considered that this was a reasonable course despite some differences between the trials. Table 5 below shows the effects for which pooled data from the TRANSFORMS and FREEDOMS trials showed a statistically significant difference between fingolimod 0.5 mg and Avonex 30 mcg, whilst table 6 shows those for which the pooled data from the trials showed a statistically significant difference between fingolimod and placebo.

As can be seen from tables 5 and 6 there were few consistent patterns of adverse events. While fingolimod was associated with significantly more influenza-type illness than placebo, the incidence was still significantly lower than was the case in the Avonex arm of TRANSFORMS. Patients treated with fingolimod also showed higher incidences of raised alanine aminotransferase (ALT), gamma-glutamyl transferase (GGT) and hepatic enzymes than those in either the Avonex or the placebo groups.

Table 5: Adverse events for which there was a statistically significant difference between the pooled fingolimod arms and the Avonex arm of the TRANSFORMS trial.

Adverse event	Fingolimod 0.5 mg (n = 854): N (%)	Avonex 30 mcg (n = 431): N(%)	RR fingolimod versus Avonex (95% CI)
Upper respiratory tract infection	86 (10.1)	27 (6.3)	1.61 (1.06 to 2.44)
Dyspnoea	36 (4.2)	7 (1.6)	2.60 (1.16 to 5.78)
Hypercholesterolaemia	24 (2.8)	3 (0.7)	4.04 (1.22 to 13.33)
Vertigo	23 (2.7)	3 (0.7)	3.87 (1.17 to 12.81)
Diarrhoea	67 (7.8)	21 (4.9)	1.61 (1.00 to 2.59)
Pyrexia	24 (2.8)	77 (17.9)	0.16 (0.10 to 0.25)
Influenza-type illness	21 (2.5)	159 (36.9)	0.07 (0.04 to 0.10)
ALT increased	61 (7.1)	8 (1.9)	3.85 (1.86 to 7.97)
GGT increased	28 (3.3)	1 (0.2)	14.13 (1.93 to 103.51)
Hepatic enzyme increased	30 (3.5)	3 (0.7)	5.05 (1.55 to 16.44)

Table 6: Adverse events for which there was a statistically significant difference between the pooled fingolimod arms and the placebo arm of the FREEDOMS trial.

Adverse event	Fingolimod 0.5 mg (n = 854): N (%)	Placebo (n = 418): N (%)	RR fingolimod versus placebo (95% CI)
Upper respiratory tract infection	86 (10.1)	58 (13.9)	0.73 (0.53 to 0.99)
Migraine	24 (2.8)	3 (0.7)	3.92 (1.19 to 12.93)
Influenza-type illness	21 (2.5)	2 (0.5)	5.14 (1.21 to 21.81)
Alanine aminotransferase (ALT) increased	61 (7.1)	11 (2.6)	2.71 (1.44 to 5.10)
Gamma-glutamyl transferase (GGT) increased	28 (3.3)	3 (0.7)	4.57 (1.40 to 14.94)
Hepatic enzyme increased	30 (3.5)	1 (0.2)	14.68 (2.01 to 107.30)
Weight increased	12 (1.4)	18 (4.3)	0.33 (0.16 to 0.67)

Results of MTC

The MTC assessed the following outcomes: ARR, disability progression and discontinuation due to adverse events for the following interventions: fingolimod 0.5 mg, natalizumab, interferon beta-1a (Avonex, Rebif), interferon beta-1b (Betaferon (50 mcg/250 mcg)), glatiramer acetate (Copaxone), and placebo. These analyses indicated that, for disability progression (manufacturer's submission table 34), the best performing treatment was [redacted], followed by [redacted] and then [redacted], but that the only statistically significant benefits were observed for [redacted] versus [redacted] (RR [redacted] 95% CI [redacted] to [redacted]) and [redacted] versus [redacted] (RR [redacted], 95% CI [redacted] to [redacted]).

For ARR (manufacturer's submission, table 35) the analyses indicated that, again, the best performing treatments were [redacted], followed by [redacted], followed in this case by [redacted]. All active treatments were [redacted]. Amongst active treatments [redacted], was statistically significantly superior to [redacted]. The comparison of [redacted] versus [redacted] produced an RR [redacted].

Finally the analyses of treatment discontinuation (manufacturer's submission, table 36) predictably indicated that [redacted]. Of these the best performing treatment was [redacted], followed by [redacted] and [redacted]. There were [redacted]; [redacted] was however statistically significantly superior to [redacted] and to [redacted].

The manufacturer's submission explored several sources of potential heterogeneity (manufacturer's submission table 39) and found that there were no covariates which were statistically significant for treatment discontinuation due to adverse effects. For ARR both baseline EDSS (p = 0.037) and publication year (p = 0.002) were statistically significant, whilst for disability progression age (p = 0.037) and timepoint of analysis (p 0.048) were statistically significant.

4.2.2 Critique of submitted evidence syntheses

The TRANSFORMS and FREEDOMS trials were well-conducted phase III trials which were appropriately powered to assess the primary outcome of ARR in each case;[1, 2]

TRANSFORMS had a duration of 12 months and FREEDOMS had a duration of 24 months. Other relevant outcomes including disability progression were also assessed in each trial.

However, in each case the trial population was broader (patients with RRMS) than the patient groups for which CHMP indicated approval. The submission therefore rested on post-hoc identification of sub-groups which approximated to those approved groups and, in particular, on an approximation to population 1b which the manufacturer selected as the base case. These subgroups showed a considerable degree of overlap. The ERG was particularly concerned by the fact that population 1b contained significant numbers of patients who met the criteria for population 2. These patients (population 1b and 2) would be eligible for treatment with natalizumab; therefore Avonex (and also BSC) cannot be considered the most appropriate comparator in this RES population.

The only synthesis of the two head-to-head trials consisted of a pooling of the fingolimod 0.5 mg arms for the assessment of adverse events relative to the placebo arm of FREEDOMS and the Avonex arm of TRANSFORMS.

An MTC of 18 trials attempted to provide evidence of fingolimod's efficacy on key outcomes of ARR, disability progression and treatment discontinuation due to adverse events. These included trials (with the exception of the AFFIRM trial which compared natalizumab versus placebo)[3] had populations who met criteria for RRMS but not necessarily for any of the populations defined by CHMP. The submission did not attempt the post-hoc identification of subgroups within these trials. There was also, as the submission noted, very considerable clinical heterogeneity between the trials with respect to permitted and actual prior use of DMT, duration, and criteria used to define disability progression. An exploration of covariates did not indicate any variables with consistent statistically significant effects on all treatment endpoints. However, two covariates with statistical significance were identified for each of ARR (baseline EDSS and publication year) and disability progression (age and timepoint of analysis).

As a consequence of the heterogeneity, and the fact that it was based on general RRMS populations (with the exception of the AFFIRM trial[3]), the MTC was not subsequently used to inform the economic model. In place of this, an indirect comparison was employed to provide an estimate of the relative efficacy of Avonex and placebo for the economic model. No details of this analysis were presented in the clinical effectiveness section of the manufacturer's submission; this is discussed in detail in section 5 below.

4.2.3 Summary

The submission rested on evidence drawn from two good quality RCTs which were adequately powered for the comparisons assessed in the whole trial populations.[1, 2] The TRANSFORMS trial (N = 1272) assessed fingolimod 0.5 mg in a head to head comparison with Avonex 30 mcg,[2] while the FREEDOMS trial (N = 1292) compared fingolimod at this dose to placebo (both trials also included fingolimod 1.25 mcg arms). However, in both trials the populations enrolled were those who met the criteria for RRMS. When the CHMP issued its positive opinion for fingolimod in specific subgroups of the RRMS population the manufacturer defined post hoc subgroups in both the FREEDOMS and TRANSFORMS trials which approximated to those indicated populations. These subgroups were considered by the ERG's clinical advisor to represent reasonable approximations to the indicated populations.

The manufacturer's submission identified population 1b as the base case. This constituted 43.6% of the population of the TRANSFORMS trial and only 19.7% of the FREEDOMS trial. There was, however, considerable overlap between the identified subgroups, which was considered particularly problematic in the case of patients who met criteria for population 2 in addition to population 1b, since the appropriate comparator for these patients would be natalizumab rather than a beta-interferon or BSC. This was addressed by the provision of data for population 1b but not 2. If the patients who also meet criteria for population 2 are excluded then the number of eligible patients in the fingolimod and comparator arms proportions is further reduced. Clearly, therefore, the initial power calculations do not give a good indication of the trials' ability to assess the comparison in the indicated population.

The other major problem with the evaluation of fingolimod in the manufacturer's submission lies in the comparator; the only head to head comparison is with Avonex. The primary problem with comparing fingolimod to any beta-interferon in population 1b is that this population has, by definition, failed to respond to prior DMT which, in almost all cases, consisted of beta-interferon therapy. In addition there is conflicting evidence as to the efficacy of Avonex relative to other formulations of beta-interferon covered by the NHS RSS. The selection of population 1b with a comparator of Avonex as the base case for the submission therefore appears highly likely to be non conservative.

The MTC presented in the clinical effectiveness sections related to the RRMS populations generally, contained a significant level of clinical heterogeneity, and was not used to inform the economic model, which also did not include a comparison with BSC. Therefore the non-ideal comparator Avonex constitutes the only benchmark for the relative efficacy of

fingolimod. Given that the indirect comparison presented for population 1b but not 2 indicated Avonex to be less cost-effective that placebo, while Avonex was dominated or extendedly dominated for both populations 1b and 1b but not 2, the appropriateness of this is clearly open to question.

5 ECONOMIC EVALUATION

This section focuses on the economic evidence submitted by the manufacturer and the additional information provided following ERG points of clarification. The submission was subject to a critical review on the basis of the manufacturer's report and by direct examination of the electronic version of the economic model. The critical appraisal was conducted with the aid of a checklist to assess the quality of economic evaluations and a narrative review to highlight key assumptions and possible limitations (Appendix 1). Section 6 presents additional work undertaken by the ERG to address any remaining uncertainties.

The manufacturer's initial economic submission included:

1. A description of the systematic search strategy used to identify existing cost-effectiveness studies for fingolimod in the treatment of relapse remitting multiple sclerosis (RRMS) (Manufacturer's Submission, Section 6.1) with full details in a separate Appendix (Manufacturer's Submission, Appendix 10).

2. A report on the *de novo* economic evaluation conducted by the manufacturer. The report described the technology; comparators and patient population; the categories of resource use costed; the resource use and unit cost assumptions and sources; the assumptions and sources of evidence used to assess quality of life; the base-case cost-effectiveness results; and sensitivity analysis (Manufacturer's Submission, Section 6.2).

3. An Excel-based model comprising the manufacturer's electronic economic model. The ERG has noted that the Excel-based model allows the user to produce the results of scenarios that are not presented or discussed within the main submission. As there was no clear reporting of the underlying assumptions of these additional analyses, the ERG has chosen to disregard these scenarios from the critical review.

In response to the request for clarification made by the ERG, the manufacturer further submitted:

4. A descriptive reply to the ERG's points of clarifications.

5. A brief report on the application of the *de novo* economic model to a subgroup of patients. This report described the clinical evidence available for this subgroup, and modifications made to the economic model input data as a consequence of this evidence.

6. An Excel-based model of the analysis undertaken for this subgroup.

The manufacturer undertook a systematic review of cost-effectiveness studies including fingolimod for the treatment of adult patients with RRMS. No cost-effectiveness evidence was found in the literature, thus a de novo economic model was developed. The model presented evaluates the cost-effectiveness of fingolimod compared with Avonex. No other comparators were considered. The original submission focused on a sub-population of adults with highly active RRMS, with high disease activity despite treatment with a beta-interferon and with an unchanged or increased relapse rate or on-going severe relapses as compared with the previous year (population 1b, as defined previously in Section 3.1). Results were not provided for any other sub-groups in the original submission.

The cost-effectiveness evaluation used a decision model, designed as a Markov model, to model disease progression using 21 health states representing different degrees of disease severity (by tracking EDSS scores whilst in RRMS and after conversion to SPMS) and death. Disability progression and conversion to SPMS were assumed irreversible. The model also accounted for relapses, adverse events, withdrawal and death. Although the occurrence of relapses did not influence the way in which progression was modelled to occur, relapse was modelled to depend on EDSS score. After withdrawing from fingolimod or Avonex patients were assumed to receive BSC.

The perspective of the analysis of costs was that of the NHS and PSS. Costs were separated into disease costs, administration and monitoring costs and drug acquisition costs. QALYs were used as the measure of outcomes. Both patient and caregiver utility were accounted for and varied by disease severity. Utility adjustments were also applied to account for relapses and adverse events. Treatment with fingolimod or Avonex was assumed to be provided only to RRMS patients with an EDSS score of between 0.0 and 6.0. Patients were modelled to continue to receive these treatments until the treatment was either withdrawn (due to adverse events, disease progression to an EDSS score of above 6 or conversion to SPMS) or a patient died. A 50 year time horizon was used in the model to ‘sufficiently capture differences in costs and outcomes’; previous appraisals in MS assessed by NICE adopted time-horizons lower than or equal to 20 years (TA 32 and TA 127).[4, 5] Both costs and benefits were discounted at 3.5%.

Natural history data were derived from external observational data sets and were used to inform the key events in the model. Natural history was assumed to represent the course of disease under BSC; although the cost-effectiveness results of this treatment strategy were not presented in the submission these were available in the model spreadsheet. External

observational datasets were used despite data also being available from the clinical trials (FREEDOMS and TRANSFORMS) evaluating the use of fingolimod, Avonex and placebo. The treatment effects for fingolimod and Avonex for disability progression and relapse were derived from these two clinical trials and were applied as relative risks to the appropriate summaries of the natural history data (assuming the comparison with placebo in the FREEDOMS trial is representative of a comparison against BSC). Although FREEDOMS and TRANSFORMS followed patients for 24 and 12 months respectively, in the model treatment effects were assumed to be sustained for as long as patients remained on treatment.

Data on mortality were derived from national mortality statistics and were adjusted for the additional risk of mortality for different EDSS states. Adjustment factors were derived from a combination of published studies (Pokorski (1997) and Sadovnik (1992)).[6, 7]

Both deterministic and probabilistic sensitivity analysis were carried out by the manufacturer to demonstrate the level of uncertainty around the model results. Despite the non-linear nature of the model, only the deterministic results were presented in the manufacturer's submission. The incremental cost-effectiveness ratio (ICER) of fingolimod relative to Avonex in population 1b was estimated to be £55,634 per QALY gained. The corresponding probabilistic estimate of the ICER (based on additional analysis carried out by the ERG maintaining the manufacturer's other assumptions for the base case) was estimated to be £69,787 per QALY gained.

In response to the ERG's request for further evidence, the manufacturer submitted additional cost-effectiveness analysis for population 1b but not 2 (based on the model described above). The model inputs modified to make predictions for this sub-population were patient characteristics and the estimates of the relative effectiveness of treatments. The ICER for fingolimod compared to Avonex was estimated to be £18,741 per QALY gained i.e. more favourable than in the previous analysis. The difference in the estimates was largely due to the revised relative efficacy estimates for Avonex suggesting that in this sub-population it was [redacted]. The relative risk of progression of Avonex vs. placebo was [redacted] in population 1b (i.e. approximately a [redacted]% [redacted] in the risk) and [redacted] (i.e. approximately a [redacted]% [redacted]) in population 1b but not 2. This additional analysis raised important questions concerning the robustness of the model and the decision to restrict comparisons to be solely against Avonex.

5.1 ERG comment and critique on manufacturer's review of cost-effectiveness evidence

The manufacturer's review was primarily aimed at the identification of previously published cost-effectiveness studies of fingolimod for the treatment of adults with RRMS. Additional aims included the identification of reviews of utility estimates in multiple sclerosis and reviews of resource use and cost estimates in multiple sclerosis. The databases searched for the cost-effectiveness section included all of those specified by NICE in the specification for manufacturer/sponsor submission of evidence; MEDLINE, MEDLINE In-Process, EMBASE, EconLIT and NHS EED. Searches were also carried out of the Cochrane Library, including:

- the Cochrane Database of Systematic Reviews
- the Cochrane Central Register of Controlled Trials
- the Database of Abstracts of Reviews of Effects
- the Health Technologies Assessment database.

A group of organisation's websites were also searched to identify conference abstracts and unpublished studies, these were:

- the International Society for Pharmacoeconomics and Outcomes Research
- the American Academy of Neurology
- the Americas Committee for Treatment and Research in Multiple Sclerosis
- the European Committee for Treatment and Research in Multiple Sclerosis
- the European Charcot Foundation

In addition the NICE website was searched to identify any relevant Health Technology Assessment reports as well as a search of the bibliographies of the 'seminal' papers.

The submission gave detailed descriptions of the search strategies used to obtain papers from the databases and met NICE requirements. It included the specific databases searched; the service providers used; the dates when searches were conducted; the date spans of the searches; and the complete strategies used. The strategies aimed to retrieve all research relating to multiple sclerosis, treatments for multiple sclerosis, utility studies and economic evaluation. The terms used for each search facet were appropriate. Truncation and wildcards were used appropriately.

The ERG considers the search strategy for cost-effectiveness (MS, Appendix 10, section 9.10) to be appropriate. The ERG believes that the searches were unlikely to miss any

published studies relating to any of the three aims of the search strategy that could be potentially useful in the cost-effectiveness section of the submission.

The searches conducted by the manufacturer identified 891 unique potential records from the databases and 2,513 conference abstracts.

Of these identified references, none were determined to meet the primary objective of the review given as the performance of a systematic search to identify all existing economic evaluations of fingolimod for the treatment of adults with RRMS. Of the secondary aims of the search strategy, thirteen of the identified references met the aim of a review of utility estimates, and twelve satisfied the review of resource use and cost estimates.

As no references met the primary aim of the search, the manufacturer's de-novo model was the sole focus of the submission. The ERG would like to highlight the lack of inclusion of cost-effectiveness studies which did not include fingolimod as a treatment. It is the opinion of the ERG that in doing so the manufacturer has limited the evidence base used to inform the development of their model structure and many of the key assumptions underpinning their model.

5.2 Summary and Critique of Manufacturer's Submitted Economic Evaluation by the ERG

An overall summary of the manufacturer's approach and signposts to the relevant sections in the manufacturer's submission are reported in Table 7 below:

Table 7: Summary of the Manufacturer's economic evaluation (and signposts to manufacturer's submission)

	Approach	**Source / Justification**	**Signpost (location in manufacturer submission)**
Model	Markov cohort model that tracks disability progression (using the EDSS scale), the occurrence of relapses and the conversion from RRMS to SPMS.	The manufacturer's submission justifies that the model captures the disability associated with MS, and the potential for relapse. The structure of the model is similar to previous NICE technology appraisals in MS (TA32 [NICE, 2002] and TA127 [NICE, 2007]).[4, 5]	*Sections 2.1, 6.2.2 and 6.2.3*
States and events	The model includes a total of 21 health states: each of 10 EDSS states (aggregated to the values of 0, 1, 2, … , 9) for RRMS and SPMS, as well as	This approach to defining states and events in multiple sclerosis is consistent with previous appraisals.	*Section 6.2.3*

	death. Progression through these states is irreversible. The occurrence of relapses, adverse events and withdrawal from treatment was also allowed.		
Comparators	Although the scope by NICE listed comparators as interferon beta, glatiramer acetate and natalizumab (the latter only for people with rapidly evolving severe RRMS), the only comparator used in the submission was a specific beta-interferon product (Avonex).	The use of Avonex as the comparator is justified in the submission by the availability of data from the FREEDOMS trial. The exclusion of other comparators is justified based on the lack of head to head trials.	Sections 2.6, 5.2.5 and 5.2.6
Natural History	Based on the Markov model (discussed above), natural history comprises disability progression, relapse, conversion from RRMS to SPMS and mortality. Natural history was assumed in the submission to represent the course of disease under BSC.	Data for disability progression and conversion were derived from a large observational data set – the London Ontario dataset. Data for relapse rate is calculated using a study by Patzold and Pocklington (1982).[37] Data for mortality are derived from national mortality statistics adjusted for the additional risk of mortality for different EDSS states (Pokorski (1997) and Sadovnik (1992)).[6, 7] Natural history data was mainly derived from external observational data sets, despite data also being available from the clinical trials, FREEDOMS and TRANSFORMS. No justification for this was provided by the manufacturer	*Section 6.2.3*
Treatment effectiveness	The model assumes treatments to impact on the rate of disease progression and on occurrence of relapse. Mortality rates are assumed to be treatment independent but because mortality is linked to disability progression there will be an indirect effect of treatments on mortality.	Treatment effectiveness for fingolimod, Avonex and BSC are derived from the FREEDOMS and TRANSFORMS studies. It is unclear why unadjusted relative risks were used instead of the adjusted measures published when reporting the trials' results.	*Section 6.3*
Adverse events	The most severe adverse events specific to fingolimod as well as those associated with the comparator were considered. Estimated decreases in utility per SAE were applied. Costs per adverse event were also considered. Adverse events were assumed one of the main causes of treatment discontinuation (in addition to mortality and becoming treatment ineligible).	Risks of adverse events associated with fingolimod and Avonex were derived from the FREEDOMS and TRANSFORMS studies. Utility decrements associated with Avonex were derived from a study by Prosser et al. (2003).[38] Utility decrements for fingolimod were derived from a wide range of studies (Manufacturer's Submission, Section 6.4.9, Table 63). Costs per adverse event were derived using the National Schedule of	*Sections 5.9, 6.4.9and 6.5.7*

		Reference Costs 2009-10.[39]	
Health related quality of life	Separate utility values were assigned to each of the EDSS states as well as the effect of relapses and 'other' factors (years since diagnosis and gender). Utility of caregivers was also considered as a variable dependent on the patient's EDSS state. In addition, disutility from treatment was considered for Avonex only. Fingolimod was assumed to have no treatment disutility as an oral drug.	The utility values were derived from the published literature, and the disutility data from previous NICE multiple sclerosis submissions. Treatment disutilities for Avonex (derived from a study by Prosser et al. 2003) were applied on the basis that this treatment is administered as an intra-muscular injection. [38]	*Sections 6.4.5 to 6.4.9*
Resource utilisation and costs	The following cost categories were considered in the manufacturer analyses: drug acquisition costs, drug administration costs, duration of treatment, supportive care costs and adverse event costs.	The data sources used included UK reference costs, published literature and clinical expert opinion. The unit cost of drugs was based on NHS list prices (BNF 60).[40] The prices of the comparator treatments were taken once any risk sharing scheme had been applied.	*Section 6.5*
Discount rates	A 3.5% discount rate was employed for both costs and health benefits.	In accordance with the NICE reference case approach.	*Section 6.5.1*
Population and Subgroups	The main population considered (active RRMS 'non-responder' & EDSS 0 – 6.0) is itself a subgroup from the main trials. Further subgroups are identified (e.g. RES), however, no specific subgroup analysis was undertaken in the manufacturer's submission. In response to the ERG's points for clarification, a separate analysis was undertaken on the non-RES subgroup of the non-responder population – population 1b but not 2	The definition of the population in the submission constitutes only part of the population approved for licensing purposes (European Medicines Agency's, EMA). The reason given is that this represents the subgroup with the most available data, as well as representing the area of greatest unmet need and highest clinical efficacy. The submission states the fact that data are available only for a small sample of patients meant an analysis of the RES sub group was not possible.	*Section 6.2.1 and 2.4*
Sensitivity analysis	Scenario analysis and probabilistic sensitivity analysis (PSA) were undertaken.	Structural and deterministic sensitivity analyses are presented. A scatter plot and a cost-effectiveness acceptability plane are presented for the base case.	Sections 6.7.7 to 6.7.10

5.2.1 NICE Reference Case Checklist

Table 8 summarises the economic submission using a checklist based on NICE's reference case and other methodological recommendations, and the ERG's comments on whether the de-novo evaluation meets the requirements of these recommendations.

Table 8: NICE reference case checklist

Elements of the economic evaluation	***Reference Case***	***Included in submission***	***Comment on whether de-novo evaluation meets requirements of NICE reference case***
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	No	The cost-effectiveness analysis is presented using Avonex (a specific beta-interferon product) as the only comparator. Other beta-interferons used in clinical practice are excluded; the use of Avonex as the single specified comparator is not adequately justified. The NICE scope further specified glatiramer acetate as a comparator. The manufacturer justifies having disregarded this comparator on the basis of there being a lack of robust evidence. In addition the NICE scope specified comparison with natalizumab dismissed in the submission as being inapplicable to the sub-population of interest.
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Perspective on costs	NHS and PSS	Yes	NHS and PSS costs have been taken into account
Perspective on outcomes	All health effects on individuals	Yes	QALY benefits associated with disability progression of individuals and their caregivers were considered.
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	Time horizon analysed was 50 years
Synthesis of evidence on outcomes	Systematic review	Yes/No	Although a systematic review on treatment effectiveness measures and a mixed treatment comparisons (MTC) were conducted, these results were not used to inform cost-effectiveness. Instead a separate, indirect analysis relying on specific trials, FREEDOMS and TRANSFORMS, was used. The search for evidence on many other input parameters did not appear to be based on a systematic process.
Measure of health effects	QALYs	Yes	
Source of data for measurement of HRQL	Reported directly by patients and/or caregivers	Yes	Evidence from the published literature was used to assign a HRQL value to each EDSS state based on the EQ-5D. Disutility of adverse events was also derived from the published literature.

Source of preference data for valuation of changes in HRQL	Representative sample of the public	Yes	Utility values were based on EQ-5D estimates reflecting public preferences.
Discount rate	Annual rate of 3.5% on both costs and health effects	Yes	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	Probabilistic sensitivity analysis was conducted as well as deterministic and structural sensitivity analyses. Results for the probabilistic sensitivity analysis were presented graphically for the comparison of fingolimod versus Avonex.

Abbreviations: HRQL, health related QoL; NHS, National Health Service; PSS, personal social services; QALYs, quality-adjusted life years

5.2.2 Population

The manufacturer's submission considered the sub-population of RRMS patients with high disease activity despite treatment with a beta-interferon, the definition of a "non-responder" in this sub-group was defined as patients who have an unchanged or increased relapse rate or ongoing severe relapses, as compared to the previous year (population 1b as defined in Section 3.1). The justification provided for the selection of this sub-population is that it represent the largest subgroup of the trials used (FREEDOMS and TRANSFORMS) in addition to being the subgroup with 'greatest clinical unmet need' (manufacturer's submission).

As was highlighted in Section 3 of this report, the population analysed (population 1b) is likely to be heterogeneous and include patients with RES RRMS (population 2). Patients with RES RRMS are a very different population, for whom natalizumab rather than Avonex is the appropriate comparator (see Section 5.2.3 for more detail). As part of the ERG's requests for clarifications, analysis of the effectiveness and cost-effectiveness of the interventions in the sub-populations excluded in the main submission were requested. The ERG considers the sub-populations (and relevant comparators for consideration in each of these) to differ sufficiently such that cost-effectiveness should be considered separately for each subpopulation.

Baseline characteristics of the analysed sub-population (population 1b) were provided by the manufacturer for the TRANSFORMS study (Manufacturer's Submission, Section 5.10.4, Table 46), as well as for the FREEDOMS and TRANSFORMS pooled analysis (as used further in the model, see Table 99 below.

Table 9: Baseline patient characteristics across studies (Manufacturer's Submission, Section 6.2.1, Table 48)

Characteristic	FREEDOMS and TRANSFORMS pooled analysis of non-responder subgroup data
Age, years	37.3 (mean)
Female-to-male ratio	2.3:1
Time since first diagnosis, years	6.25 (mean)
Cohort size, n	603
Type of MS	RRMS: 100%

The patients' initial distribution by EDSS states is also considered by the manufacturer. Figure 1 below shows the patient distribution of the non-responder subgroup from the FREEDOMS and TRANSFORMS trials (pooled) as used in the model. It is interesting to note that more than 5% of the patients have EDSS scores of 0 and more than half the patient population has EDSS scores of 2 or less. It is difficult to reconcile the low EDSS scores in the trial with the definition of the modelled subgroup (a severe population, not responding to conventional treatment). In the model no patients are assumed to start with a baseline EDSS score of 6 or above, this is due to the fact that both the FREEDOMS and the TRANSFORMS trials that inform the model only recruited patients with a baseline EDSS of up to 5.5.

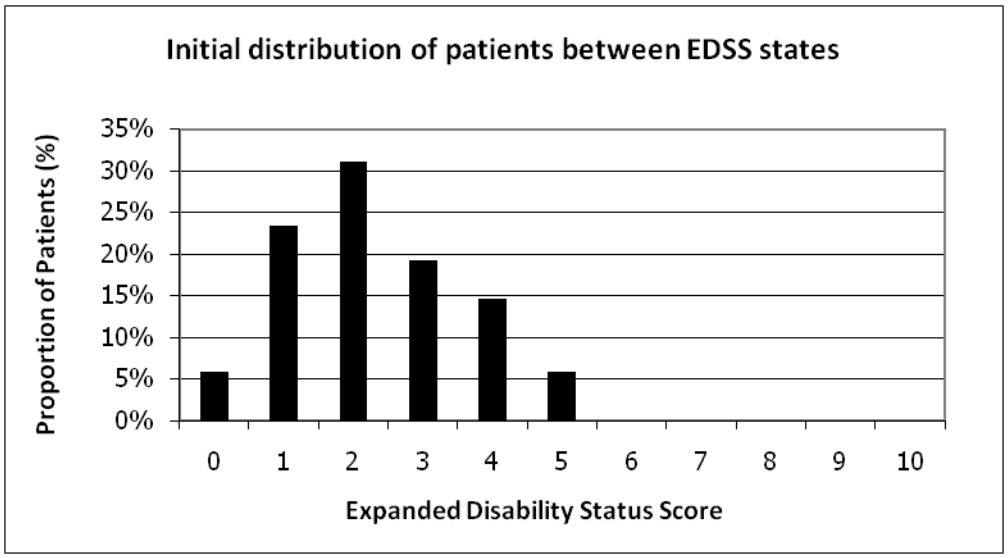

Figure 1: Distribution of patients across EDSS States, as used to inform cost-effectiveness (pooled data from FREEDOMS and TRANSFORMS)

The submission presents data on the distribution of patients across EDSS states for several different MS studies (shown in Figure 2). The figure confirms that the subgroup of patients analysed from the FREEDOMS and TRANSFORMS trials have lower EDSS scores than those seen in the other studies. This suggests that the trial samples (used further in the model) may not be representative of the non-responder population within routine clinical practice. This has not been adequately discussed or addressed in the manufacturer's submission.

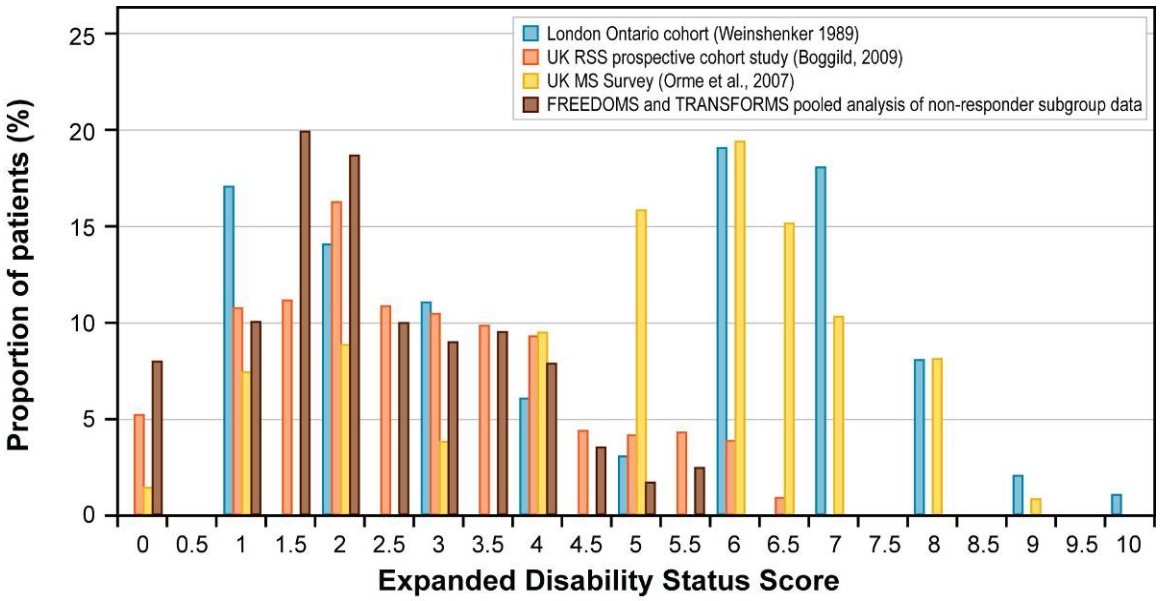

Figure 2: Distribution of Patients across EDSS States (Manufacturer's Submission, Section 6.2.1, Figure 10)

In accordance with the guidelines of the Association of British Neurologists (ABN 2009) patients in the model are only eligible for new treatment if they are able to walk independently.[41] The manufacturer states that this corresponds to an EDSS score of 6 or less, where an EDSS score of 6.0 is defined as 'Intermittent or unilateral constant assistance (cane, crutch, brace) required to walk about 100 meters with or without resting'. The use of EDSS based bounds to specify treatment ranges, and the specific ranges used, are broadly consistent with previous published models in MS.

5.2.3 Interventions and Comparators

The intervention considered in the manufacturer's submission is the licensed dose of fingolimod (0.5 mg taken once daily). The FREEDOMS and TRANSFORMS studies also evaluated a higher dose of fingolimod (1.25mg), this unlicensed dose was not included in the cost-effectiveness analysis. The manufacture only provided analysis for one comparator, Avonex (an interferon-beta 1a). The justification given for the sole use of Avonex as a comparator in the analysis was the availability of relevant data: this treatment was used as a comparator to fingolimod in the clinical TRANSFORMS study. As was highlighted in Section 3.3 of this report, existing evidence suggests other beta-interferons may be more effective than Avonex. The use of Avonex as a comparator, where more effective beta-interferons exist, may have resulted in an overestimation of the cost-effectiveness of fingolimod.

In the analysis informing cost-effectiveness, patients were assumed to continue to receive fingolimod until the treatment was withdrawn (i.e. due to adverse events, disease progression to an EDSS score of above 6, or conversion to SPMS) or a patient died. Patients for whom the treatment was withdrawn were then assumed to receive BSC alone. A similar approach was taken to model the comparator (Avonex).

In the manufacturer's initial submission, data was provided on the market share of each of the RRMS treatments licensed for use in England and Wales in the second quarter of 2010 (Manufacturer's Submission, Section 2.6, Table A8).The data show Avonex as the third largest of the RRMS treatments with 17.5% market share. When asked in the points of clarification to justify the use of Avonex as the sole comparator to fingolimod, the manufacturer provided new data (given below as Table 10) showing Avonex having the largest market share according to Prescription Pricing Authority data. The manufacturer also commented that a comparison to any other DMT would require the use of indirect comparison methods. The ERG would like to note that indirect comparison methods were used extensively in the submitted analysis, both to derive relative effectiveness inputs (from the FREEDOMS and TRANSFORMS trials) and to inform the cost-effectiveness model.

Table 10: Market share of DMTs prescribed for RRMS from Jan 2008 to June 2010 (Manufacturer's Clarifications)

Therapy	Patient share (%)
Interferon-beta-1a (Avonex)	33.8
Interferon-beta-1a (Rebif)	28.4
Glatiramer acetate (Copaxone)	21.3
Interferon-beta-1b (Betaferon and Extavia)	16.5

The comparators listed in the NICE scope for this appraisal are:

- Interferon beta (multiple interferon beta treatments exist)
- Glatiramer acetate
- Optimised standard care with no disease modifying treatment (Best Supportive Care (BSC))

In addition, for people with RES RRMS (population 2, as defined in section 5.2.2)

- Natalizumab

In addition to meeting the requirements of the NICE scope, the ERG deems a comparison against BSC to be important since the sub-population considered in this analysis is one where patients have failed to respond to a previous course of DMTs. The cost-effectiveness of continued use of beta-interferon (or switching to an alternative product) in this subpopulation has not been evaluated in previous NICE appraisals and hence it should not be assumed that continued use of a beta-interferon is, in itself, cost-effective. Comparisons with BSC were not presented in the submission.

The manufacturer also did not compare natalizumab and fingolimod in population 2. The reason given for this was that different definitions of RES were used by the AFFIRM study (used in the evaluation of natalizumab) and the FREEDOMS and TRANSFORMS studies (used in the evaluation of fingolimod). The ERG feels that this analysis would have been informative even if performed under the assumption that the two RES populations are similar.

5.2.4 Model Structure

In the absence of any previous published cost-effectiveness studies for fingolimod the manufacturer presented a *de novo* economic evaluation based on a decision model that

described the natural history of patients with RRMS. The manufacturer provided an overview of the model structure reproduced below in Figure 3. No interpretation was provided to explain the shaded area surrounding the majority of the model and some of the lower EDSS states.

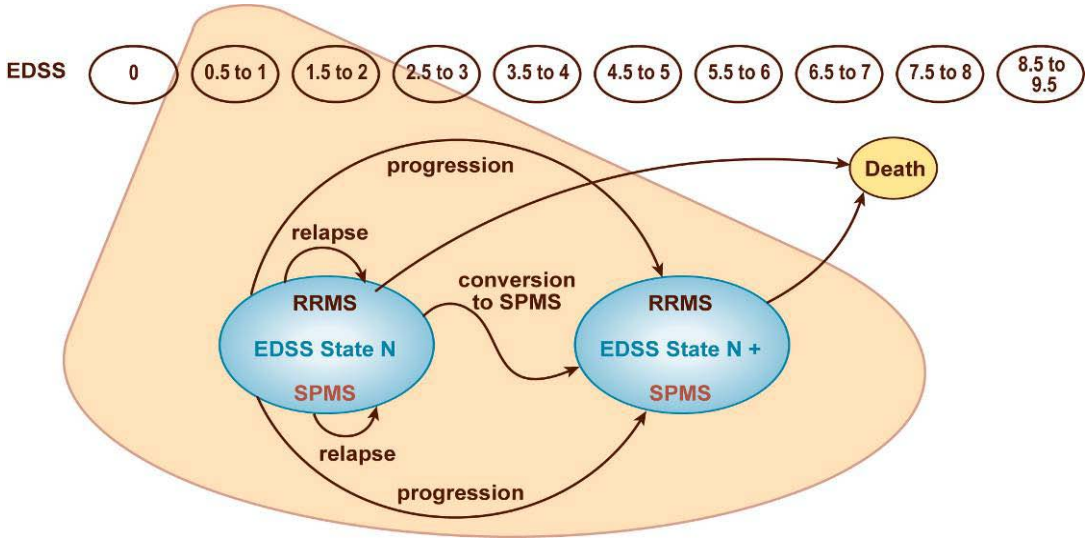

Figure 3: Model Structure (Manufacturer's Submission, Section 6.2.2, Figure 11)

EDSS, Expanded Disability Status Scale; RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis.
Source: Biogen Idec UK and Elan Pharma International, 2007.

The model is structurally similar to the models used in previous NICE submissions in MS (for example the ScHARR model). It uses a Markov structure and focuses on five dimensions of the disease:

A. Disability progression
B. Conversion from RRMS to SPMS
C. Relapse
D. Mortality
E. Treatment discontinuation & Adverse Events

Each of these will be discussed in turn in this section. The key assumptions associated with each dimension are given in Table 11 at the end of this section. The data sources used to inform these dimensions will be explained in the following section on natural history

A - Disability Progression

The manufacturer chose to model disability progression in MS in a similar way as previous NICE submissions (TA32 and TA127), by assuming that both RRMS and SPMS patients experienced an underlying disability progression risk.[4, 5] The ERG would like to note that the model is driven by disability progression, rather than this progression being driven by the occurrence of disabling relapses (typical of RRMS). The ERG would also like to highlight that the modelling of RRMS and SPMS patient transitions in the same manner may not fully take account of the innate differences between the two types of MS patients.

Disability progression represents the progressive nature of MS and is present from entering the model until mortality. Disability progression was modelled as progression in EDSS. The EDSS is the standard method for quantifying the severity of MS, and represents a 20 level scoring system (between 0 and 9.5) and is derived by grouping disability into eight functional systems (FS) allowing neurologists to assign a Functional System Score (FSS) in each of these. The functional systems used are: pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, cerebral and other. The lower EDSS scores of 1.0 to 4.5 represent people who are fully ambulatory, while EDSS scores of 5.0 to 9.5 represent those who have an ambulatory impairment. An EDSS score of 10.0 refers to patients who have died due to MS related symptoms. It is worth noting that previous submissions to NICE [TA32] have highlighted that, although the EDSS scores are frequently used in MS decision models, they do have some well known limitations as a method for mapping MS disability. While an alternative method does not seem available, it is important to keep these limitations in mind.

Disability progression was defined in the manufacturer's model as an increase in EDSS score, thus in each cycle of the model an individual with RRMS can progress to a worse EDSS state. No account was thus made in the submitted model for possible regression in EDSS. It is worth noting that in the previous appraisal of natalizumab, backward transitions to improved EDSS health states (PenTAG paper, TA127) were allowed, based on data from the AFFIRM trial. However, the use and impact of this approach was not explored in detail in this appraisal.[5, 42]

The manufacturer's model reduces the 20 EDSS levels to 10, by rounding the half levels. The rounding of EDSS states in this way may impact on the assessment of the cumulative probability of sustained progression of disability resulting in an overestimation of the rate of disability progression (i.e. more rapid) as discussed in the PenTAG report on natalizumab (TA127) where the same approach was taken.[42]

B - Conversion

The model presented in the submission assumed that patients face a transition probability of conversion to SPMS for each period they are in RRMS. Once they have converted to SPMS they are assumed to be unable to revert back to RRMS. SPMS patients experience disease progression through increases in EDSS score, analogously to RRMS patients. It is unclear from the manufacturer's submission how conversion was characterised.

C - Relapse

The relapsing-remitting nature of MS was included in the model through a probability of relapse in each cycle of the model up until death. Relapse rates were modelled to depend on EDSS state, and were allowed to differ between RRMS and SPMS patients. In the manufacturer's model clinical progression was not driven by repeated relapses. The ERG's clinical understanding is that disability progression often occurs after relapse, such that a patient may never fully recover after experiencing a relapse (i.e. experiences disability progression). No account is made of this by the manufacturer.

D - Mortality

Mortality was included in the model by considering all-cause mortality. Probabilities for all-cause mortality for the general population were derived from age and gender specific mortality rates for England and Wales (ONS, 2010).[43] The probabilities were subsequently adjusted for the MS population, using the mortality ratios reported by Pokorski (1997) coupled with an assumption that MS patients with an EDSS score of 0 do not face any additional mortality risk.[6] The Pokorski ratios consist of three ratios for mild, moderate and severe MS, and are applied onto the EDSS scale as shown in Figure 4 below by the grey bars. The Pokorski ratios were subsequently used to inform a predictive function of the relationship between EDSS and MS mortality (this makes use of analysis by Sadovnick et al. (1992)).[7] The manufacturer's analysis used this predictive function (depicted in Figure 4) to derive excess mortality ratios, one for each EDSS category.

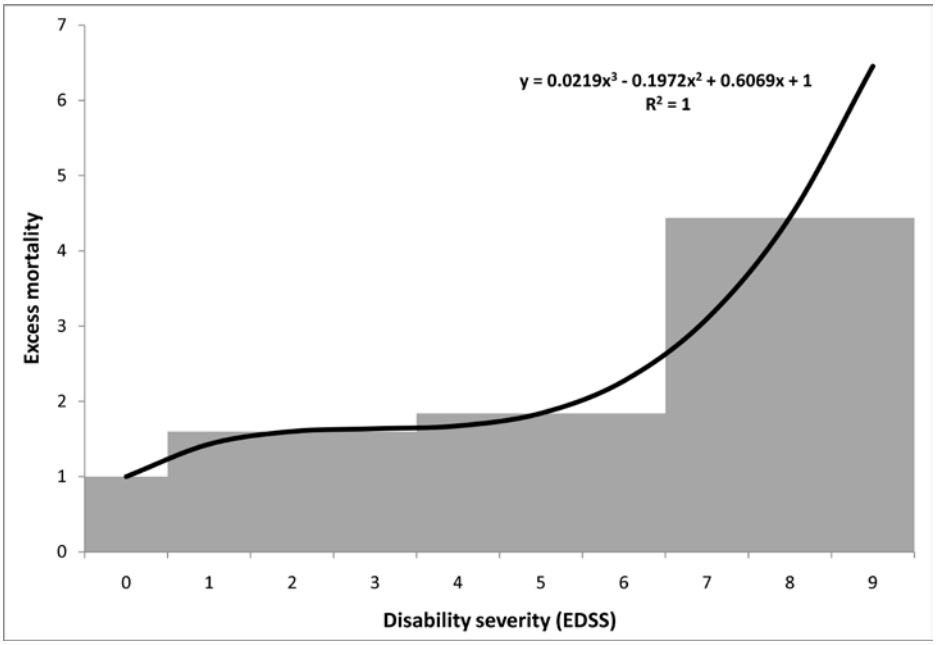

Figure 4: Effects of EDSS on mortality (Manufacturer's Submission, Section 6.2.3, Figure 13)

The manufacturer's model assumed that RRMS patients and SPMS patients who had the same EDSS score had the same mortality risk. The ERG was not provided with evidence supporting this assumption within the submission.

E - Treatment discontinuation & Adverse Events

The manufacturer outlined three situations (in addition to mortality) when patients withdrew or were withdrawn from treatment:

- When patients convert from RRMS to SPMS.
- When patients' disease progression reaches an EDSS score greater than 6. This as the EDSS upper bound for the provision of treatment is consistent with the guidelines of the Association of British Neurologists (ABN 2009).[41]
- After experiencing a serious adverse event related to the treatment.

General

The decision model takes the NHS and PSS perspective as per the NICE reference case. The manufacturer used a time horizon of 50 years in the base case, on the basis of being able to 'sufficiently capture differences in costs and outcomes'; however, other appraisals in MS assessed by NICE adopted time-horizons lower than or equal to 20 years (TA 32 and TA 127).[4, 5] The manufacturer considers the impact of varying the time horizon as part of the sensitivity analysis in the submission. The studies used to estimate relative effectiveness

(FREEDOMS and TRANSFORMS) followed patients for 24 and 12 months, respectively, and in the manufacturer's base case analysis the model extrapolates this estimate as a constant treatment effect, applied for as long as patients remained on treatment.

The manufacturer used a discount of 3.5% for both costs and health benefits in the model, as is stipulated in the NICE reference case.

Table 11: Structural assumptions of the model

Modelled parameter	Brief description and key assumptions made
Patient entry	• The model allows patients to enter the model at RRMS EDSS states 0 to 6 inclusive.
Disability progression	• Patient are able to progress within their MS type (ie RRMS or SPMS), to any EDSS state "worse" than their current. Patients are not able to regress EDSS states. • Disability progression occurs independently of the occurrence of relapses, although the reverse is not true. • Patients' future disease progression risk is independent of their previous disability progression history.
Maintaining health state	• In each period patients may remain in their current EDSS state as well as their current MS type.
Relapse	• In any period individuals can experience a relapse, even if maintaining their EDSS score. • The frequency of the occurrence of relapses depends only on EDSS state and whether the patient is RRMS/SPMS. • A patient who relapses faces the same risk of disability progression as a patient who has not.
Conversion	• RRMS patients face a probability of converting to SPMS in each period; the reverse is not allowed in the model. • If conversion occurs the individual also experiences disability progression represented by a 1 point increase in EDSS state. The exception is for RRMS patients with EDSS of 9 that convert to SPMS EDSS 9.
Discontinuation of treatment	• Discontinuation of treatment may be due to adverse events. • Conversion to SPMS leads to discontinuation. • Disability progression to an EDSS state greater than 6 leads to discontinuation.
Death	• In each period an individual faces a risk of death (adjusted for MS).
Extrapolation	• A constant treatment effect is applied for as long as patients remain on treatment throughout the 50 year time horizon of the model.

5.2.5 Natural History

The model predicts rates of progression through EDSS states, rates of conversion from RRMS to SPMS, relapse rates and mortality rates. A variety of data sources are used to predict these values for a population not receiving DMT. These values, in combination with the characteristics of the patient population used in the model, are taken to define the natural history of the disease. In the submission, natural history is assumed to represent the course of disease under BSC. The treatment effectiveness of Avonex and fingolimod are both defined in comparison to the natural history (BSC) as relative effects modifying the natural history rates of progression and relapse. These relative effects will be described in more detail in the following section on treatment effectiveness. In this section we describe and

critique how the manufacturer has determined the various elements of the natural history as used in the model.

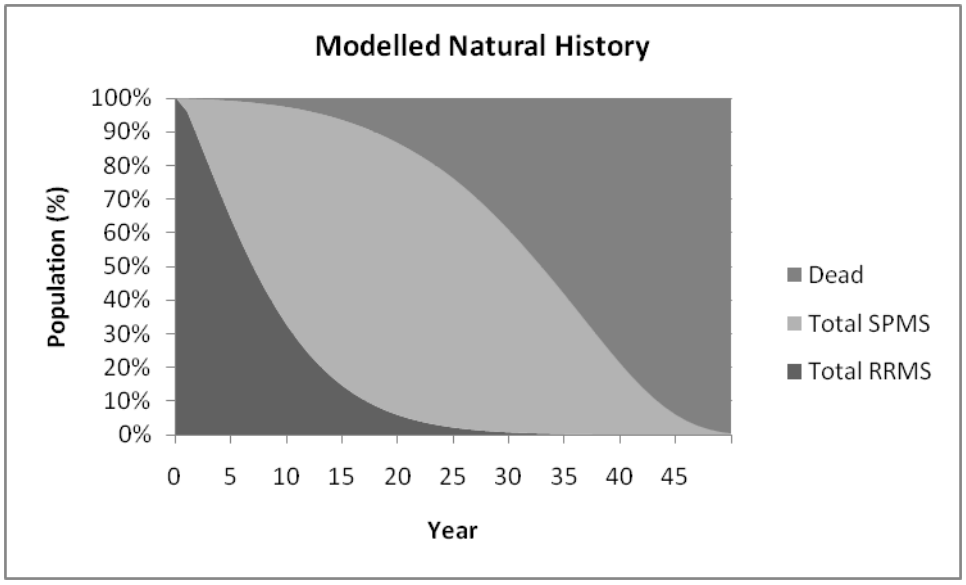

Figure 5: Natural history patient population over time (ERG analysis based on manufacturer model)

Figure 5 above illustrates how the base case model predicts patients to be distributed between RRMS, SPMS and death over the 50 year model time horizon. The model assumes all patients start in RRMS, by 25 years into the modelled period almost no patients are left in RRMS, with most having converted to SPMS and the remaining patients having died; by 50 years (the time horizon of the model) almost all patients have died. Next, we will explore each of the structural components of the model, disability progression, conversion from RRMS to SPMS, relapse and mortality.

A - Disability Progression

Disability progression was derived from a longitudinal dataset of patients with MS from London, Ontario, Canada (Weinshenker et al., 1989).[44] For this submission the manufacturer undertook a re-analysis of the dataset to provide updated transition matrices for patients who had not received DMT treatment. There was also an adjustment made for active forms of relapsing MS (this was interpreted to be for patients who reported 2 or more relapses during the first two years of MS), where patients exhibiting less progressive forms of relapsing MS were excluded. This subset of the Ontario dataset has been used to form the basis of the natural history data. The adjustments made do not seem to match the definition of the population of interest (population 1b) as described in the manufacturer's submission. Additionally, the adjustments made have not been fully described or justified making it difficult for the ERG to determine the plausibility and generalisability of the analysis.

There does not appear to have been a systematic approach to searching for evidence to describe the natural history. Despite previous models in MS having utilised the same dataset, no attempt was made in the submission to justify the use of this particular study over either the control arm of the FREEDOMS trial or any other potential external studies.

In estimating the transition matrices, the data were modelled assuming a homogeneous continuous-time Markov process (distinct from the decision model used to estimate cost effectiveness), using exponential distributions to describe transitions. The use of exponential distribution implies that the rates at which transitions between EDSS states occur were assumed to be constant over time. A proportional hazards approach was used to evaluate the effect of covariates (on patient type: active RRMS, SPMS, benign RRMS and PPMS). No attempt was made to justify or verify these assumptions. The transition matrices produced by this process are given in Table 12 (for RRMS) and Table 13 (for SPMS) reproduced below from the manufacturer's submission. These transition matrices were used as input parameters in the main decision model (the distinct homogeneous discrete time Markov process described in section 5.2.4 above). It is clear from comparing the two matrices that progression through EDSS states is predicted to accelerate upon conversion from RRMS to SPMS. An important implication of this is that any DMT impact on conversion to SPMS will also have an indirect impact on progression.

Table 12: Natural History RRMS Transition Matrix - Adjusted Ontario (Manufacturer's Submission, Section 6.2.3, Table 49) [contents of table AIC]

Table 13: Natural History SPMS Transition Matrix - Adjusted Ontario (Manufacturer's Submission, Section 6.2.3, Table 51) [contents of table AIC]

No attempt was reported in the manufacturer's submission to internally validate the transition matrices used in the model against the trial data or to externally validate these matrices

against other published natural history datasets. In order to do so, we refer to the transition matrix estimates derived using the control arm of the FREEDOMS trial extracted from the manufacturer's model (in Table 14 below). The manufacturer has not described how these data were obtained; we will assume this matrix is comparable to the ones in Table 12 and Table 13. An informal comparison of the results derived from this trial and the natural history RRMS transition matrix used in the model (Table 12) suggests that these matrices are very different. The most obvious difference being that the trial population does not seem to follow the core structural assumption enforced by the model that transitions only occur from lower to higher EDSS states.

Table 14:RRMS transition matrix FREEDOMS placebo arm - ITT analysis. (extracted from manufacturer's model)[contents of table AIC]

	0	1	2	3	4	5	6	7	8
0									
1									
2									
3									
4									
5									
6									
7									
8									

The manufacturer provided the population distribution of the 1b subpopulation from the placebo arm of the FREEDOMS trial at the start and end of the 2 year trial as part of the clarifications requested by the ERG. Using these data, we derived the population distribution at the end of the 2 year trial. We also ran the model for two years (for comparability, we used the EDSS distribution of patients in the placebo arm at the start of the FREEDOMS trial) to generate an equivalent model predicted natural history. Figure 6 below shows the initial population distribution and the trial observed and model predicted population distributions after 2 years. These results suggest that the natural history data on progression predicted by the model are not consistent with those observed in the trial, with the model predicting much faster disability progression than was observed in the trial.

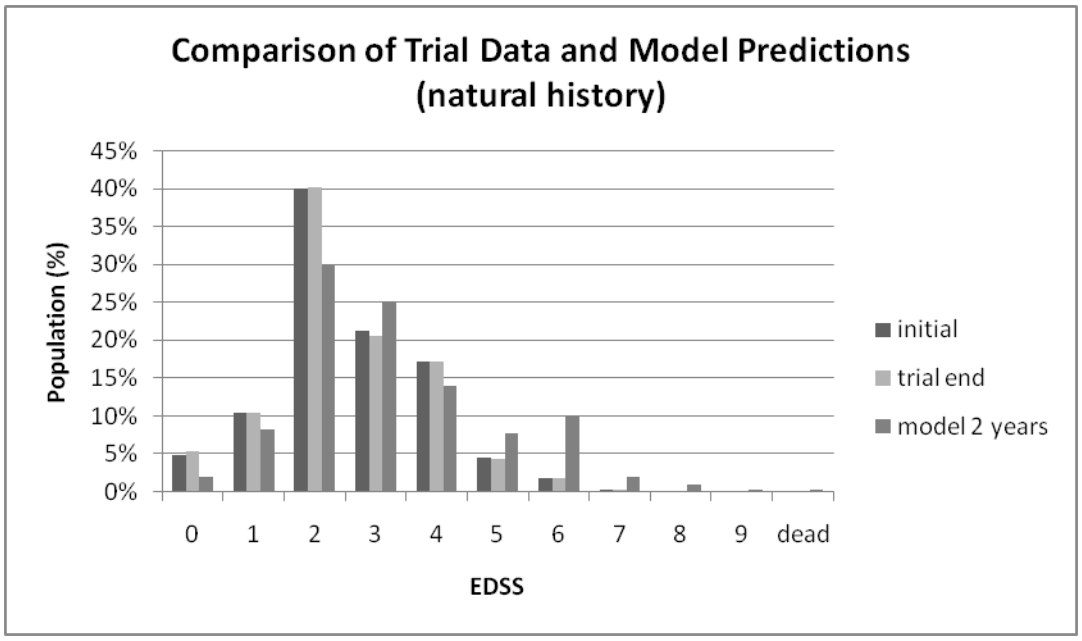

Figure 6: Comparing implications of trial and model evidence on natural history progression in RRMS (ERG analysis based on manufacturer model)

Reverting back to the pooled population distribution used in the base case results, the ERG conducted further analysis to evaluate the number of years per patient in each EDSS state split by RRMS and SPMS predicted by the model (Figure 7).

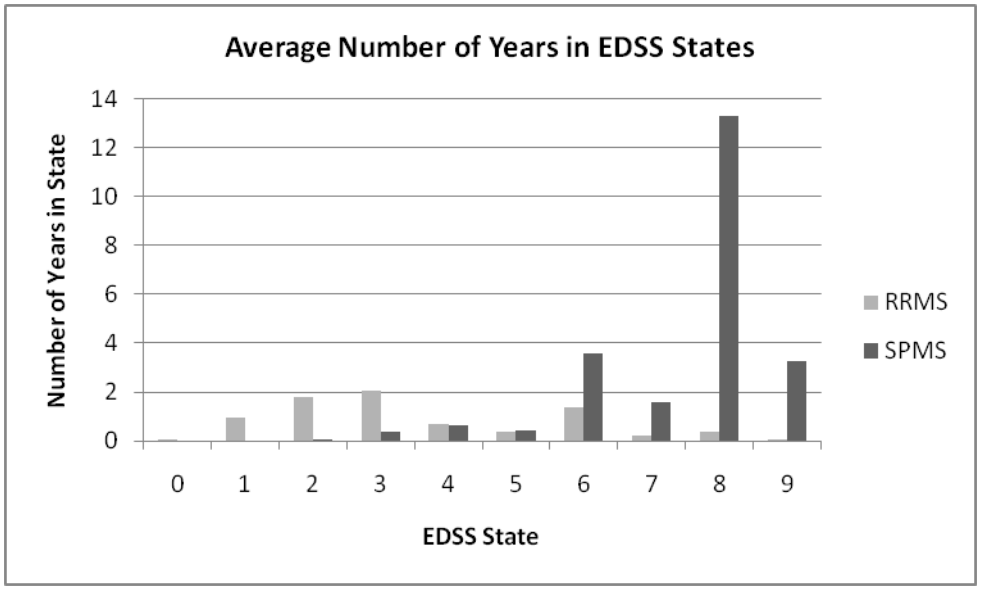

Figure 7: Natural History Years in EDSS States (ERG analysis based on manufacturer model)

Given that only patients in RRMS states between EDSS 0 and EDSS 6 are eligible for treatment, the results highlight the relatively small proportion of time the model predicts that patients spend within the treatment range.

B - Conversion

Conversion rates from RRMS to SPMS were also calculated using the London Ontario dataset. The median time to conversion observed in the data was used to specify an exponential distribution for EDSS 1 and this was used in conjunction with a Cox proportional hazards model to calculate conversion rates between EDSS 1 and all other EDSS states. Conversion probabilities for EDSS states 0 and 9 are not calculated and are assumed in the model to take values 0% and 100% respectively. A fuller description of the method used is provided in Appendix 2. Table 15 shows the conversion rates by EDSS states, as derived by the method described above, that are used in the model.

Table 15: Natural History Annual Conversion Rate by EDSS State (Manufacturer's Submission, Section 6.2.3, Table 50)

EDSS	Probability
0	[redacted]
1	[redacted]
2	[redacted]
3	[redacted]
4	[redacted]
5	[redacted]
6	[redacted]
7	[redacted]
8	[redacted]
9	[redacted]

Figure 8 below shows the proportion of patients who have converted to SPMS states over time, as predicted by the model. The model only allows for conversion in one direction i.e. from RRMS to SPMS, thus by 25 years almost all patients who are still alive have converted to SPMS.

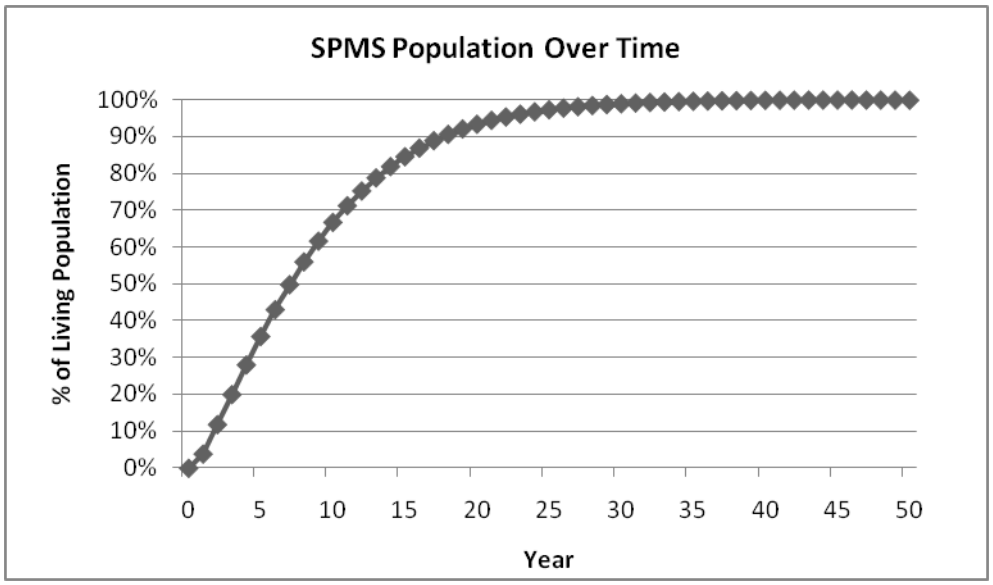

Figure 8: Natural History SPMS Population over Time (ERG analysis based on manufacturer model)

The manufacturer has not reported having conducted literature searches (systematic or not) to find evidence on conversion and has not presented any justification for using this dataset. Despite the lack of detail in reporting the analyses undertaken, the method used to calculate the conversion rates rests upon a number of unjustified assumptions such as the use of the exponential distribution. It is difficult to assess the validity of these results as there is no attempt by the manufacturer to validate the calculated SPMS conversion rates either internally against trial observations or externally against other published studies. In addition the manufacturer has not reported any conversion data from the FREEDOMS and TRANSFORMS trials for the ERG to be able to carry out its own internal validity assessment.

C - Relapse

The relapsing remitting nature of MS is incorporated in the model through the inclusion of a probability of relapse in each model cycle. These probabilities were calculated separately from the estimates for progression, hence breaking any implicit correlation between the two. The Patzold and Pocklington (1982) data on relapse rates together with the UK MS Survey (Orme et al., 2007, Tyas et al. 2007) were used to estimate the natural history of relapses by disease type (RRMS or SPMS) and EDSS state.[37, 45, 46] The data indicate an inverse correlation between the number of years since diagnosis and the annual rate of relapse, i.e. the greater the time since diagnosis the fewer relapses patients are expected to have (see Figure 9 below).

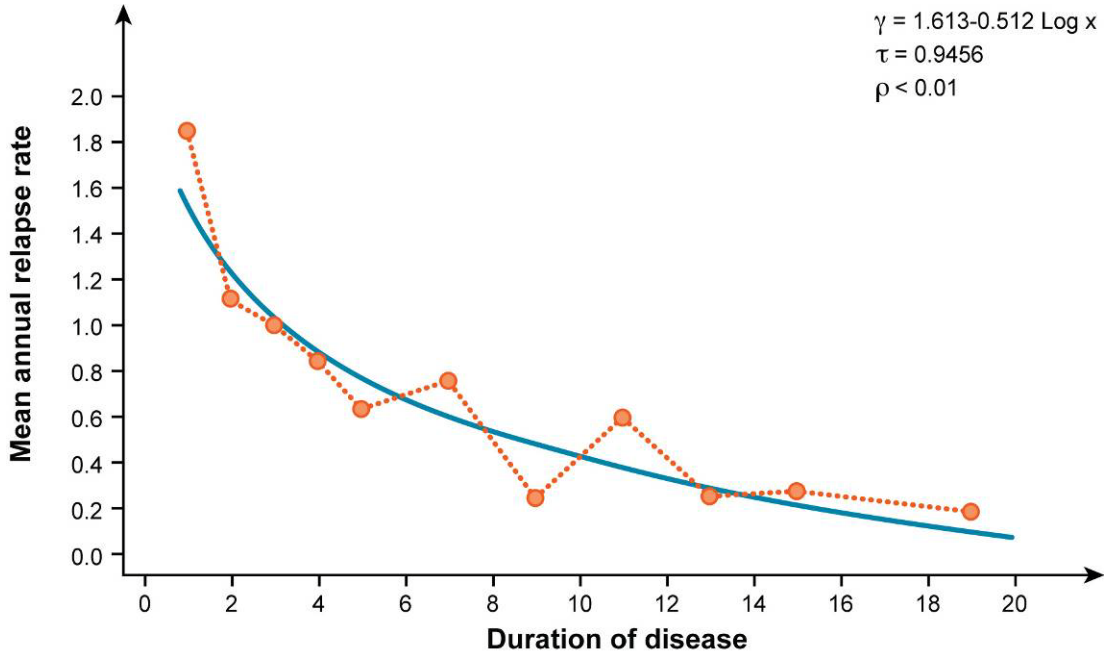

Figure 9: Negative correlation between relapse rate and duration of (Manufacturer's Submission, Section 6.2.3, Figure 12)

To estimate the relapse rates, first a matrix was constructed characterising the distribution of the patient population across each of the EDSS states by time since diagnosis using data from the UK MS Survey (Orme et al., 2007, Tyas et al. 2007).[45, 46] A similar matrix detailing relapse rates for each EDSS state by time since diagnosis was constructed from Patzold and Pocklington (1982).[37] These two matrices were then multiplied together to give a matrix estimating the number of relapses for each EDSS state by time since diagnosis. Finally, to obtain the rate of relapse for each of the EDSS states, the number of relapses in each of the EDSS states was summed across all of the times-since-diagnosis categories and then divided by the total number of patients in that EDSS state. This was done for both RRMS and SPMS states giving the results shown in Table 16 below. The ERG would like to note that the values in the table below appear highly variable and do not follow a simple directional trend.

Table 16: Natural History Annual Relapse Rates by EDSS (Manufacturer's Submission, Section 6.2.3, Table 55)

EDSS Scale	Relapse Rate	
	RRMS	SPMS
0	0.709	0
1	0.729	0
2	0.676	0.465
3	0.720	0.875
4	0.705	0.545
5	0.591	0.524
6	0.490	0.453

EDSS Scale	Relapse Rate	
	RRMS	SPMS
7	0.508	0.340
8	0.508	0.340
9	0.508	0.340

The manufacturer states that the Ontario dataset is not used here as it does not contain much relapse data and hence a combination of other studies were used in its place. There does not seem to have been any systematic search for studies on relapse rates and no justification is provided for the studies selected to calculate the natural history values. External data is used to describe both the MS patient population distribution and the relapse rates they experience in favour of trial data. No attempt to assess external validity of these relapse rates was made.

An annualised relapse rate of 0.542 was observed in the placebo arm of the relevant sub-population in the FREEDOMS trial (Manufacturer's Submission, Section 5.5.3, Table 28). The ERG established that an annualised relapse rate of 0.710 is predicted by the model for a time span equal to the trial's. There appears to be a large absolute difference between the model predicted and trial observed values. This discrepancy is not discussed by the manufacturer.

The manufacturer demonstrates a strong and very significant negative correlation between years since diagnosis and relapse rates (see Figure 9 above) and goes to great lengths to incorporate this correlation into the calculation of relapse rates by EDSS states. These relapse rates are subsequently collapsed into constant, time independent, EDSS state specific values, potentially losing much of the benefit of considering the years since diagnosis dimension of the data.

D - Mortality

Probabilities for all-cause mortality for the general population were derived from age and gender specific mortality rates for England and Wales (ONS, 2010).[43] The probabilities were adjusted for the MS population, using the mortality ratios reported by Pokorski (1997) where mortality experience of MS is categorised into mild, moderate and severe degrees of disability categories.[6] To add more granularity to these mortality ratios, analysis by Sadovnick and colleagues (1992) was used to generate an equation to predict excess mortality for individual EDSS scores.[7] The same excess mortality multipliers were applied to both RRMS and SPMS populations.

Figure 10 below shows mortality over time as predicted by the model, showing 50% of the population predicted to be dead 34 years into the modelled time horizon and almost the entire population predicted to be dead at the end of the 50 year modelled time horizon.

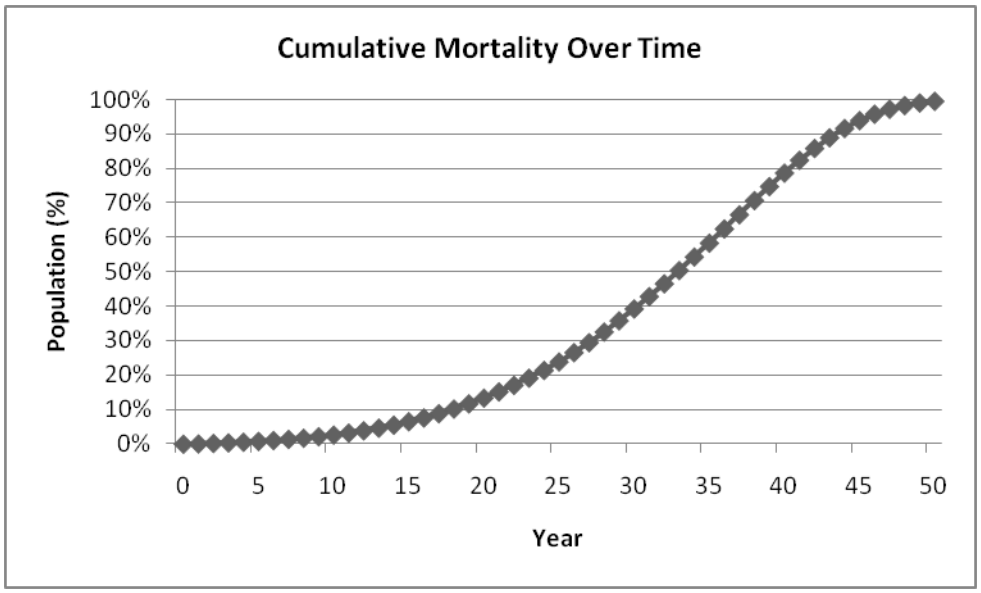

Figure 10: Natural history mortality (ERG analysis based on manufacturer model)

While the methods used to calculate mortality are similar to those used in previous NICE technology assessments for MS (TA 127) there was no evidence of a systematic approach to determine the most appropriate mortality ratios to use and no validation of these against other studies.[5]

Overall considerations

There are a number of concerns that the ERG has identified with the derivation of the natural history:

- The manufacturer does not appear to have used a systematic approach to identify and subsequently select appropriate data sources to inform the key parameters of the natural history – instead the choices of data appear to be arbitrary.
- Methods used for deriving the various elements of the natural history are not fully described and the assumptions made are not discussed or justified.
- There has been no attempt by the manufacturer to validate the predictions of the natural history either internally against the trial data or externally against other published studies.

- Where the ERG has attempted to compare natural history model predictions with results from the manufacturer's trials these seem to show substantial divergence between the two.

5.2.6 Treatment Effectiveness

Treatment effectiveness is captured in the model by applying separate relative risk estimates to the natural history estimates for both disease progression and relapse. These relative effects are calculated for fingolimod and Avonex using data from the FREEDOMS and TRANSFORMS trials. The manufacturer assumed that the relative risks remain constant while remaining on treatment, throughout the 50 year time horizon of the model. The model assumes that there is no remaining treatment effect once DMT is stopped.

The manufacturer has also conducted an MTC using a wider network of data and comparators to calculate relative treatment effects; these results however are not used in the model. This section describes how the treatment effects are applied in the model, assessing the impact of treatment on each element of the model (disability progression, conversion, relapse and mortality). The terms natural history, BSC and placebo are used interchangeably in this section reflecting the manufacturer's use of these terms in the submission.

A - Disability Progression

The effect of treatment on disease progression is modelled as a relative risk of confirmed disability progression. The model applies relative risks for fingolimod and Avonex vs. BSC to the individual natural history transitions for RRMS. Hazard ratios reported from the manufacturer's trials and quoted in the submission are not used in the model.

The relative risk for fingolimod versus BSC is derived directly from the relevant subset of the population in the FREEDOMS study. The corresponding Avonex versus BSC relative risk value is calculated using the standard adjusted indirect method (Bucher et al., 1997) based on combining the FREEDOMS study, which compared fingolimod with BSC, with the TRANSFORMS study, which directly compared fingolimod with Avonex.

The relative risks for disease progression are assumed to be constant (time independent) and are only applied while on treatment. All patients are modelled to follow a natural history path (i.e. BSC) once they fall outside the RRMS treatment range or have converted to SPMS.

Table 17 below summarises the relative risks of progression used in the model as calculated by the above methods. A full description of the calculation of these values can be found in Appendix 3.

Table 17: Relative Risk of Progression

fingolimod vs. Placebo	Avonex vs. Placebo
[redacted]	[redacted]

To illustrate the difference in predictions of progression for the two DMTs we used the manufacturer's model to calculate the average time in each EDSS state for patients modelled to be treated with fingolimod as compared to those modelled to be treated with Avonex; the results of this analysis are shown in Figure 11 below. This figure shows that those patients treated with fingolimod on average spend more time in lower EDSS states (0-2) and more time in RRMS as compared to those modelled to be treated with Avonex. The latter spend on average more years in higher EDSS states (7-9) and in SPMS. The ERG has already expressed concerns about the large proportion of the patient population in this severe non-responder RRMS subgroup modelled as starting in low EDSS states. It appears from the analysis here that much of the differences in progression occur in these low EDSS states (0-2). We will return to the issue of exploring the impact of using different initial patient population distributions across EDSS states in the model in section 6.

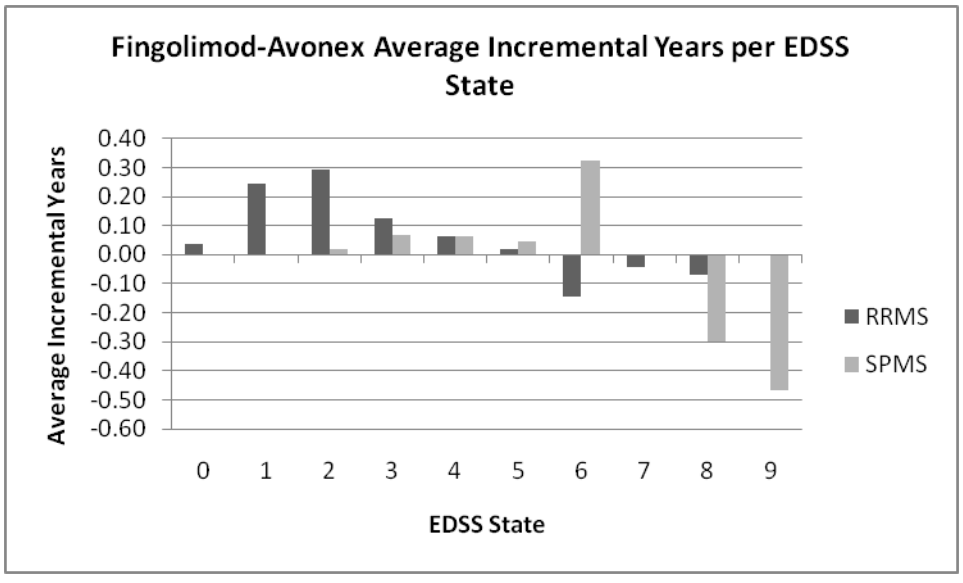

Figure 11: Incremental Years in EDSS States (ERG analysis based on manufacturer model)

As part of the clarification requested by the ERG the manufacturer provided the distribution of patients across EDSS states at the start and end of the fingolimod arm of the FREEDOMS trial. Using this same initial trial population distribution the ERG ran the model for two years

to generate an equivalent model predicted population distribution for patients treated with fingolimod (see Section 5.2.5 for the equivalent analysis on predictions over BSC). Figure 12 below shows the initial population distribution, the end of trial population distribution and the model predicted population distribution after 2 years. This analysis suggests that the predicted progression results for patients treated with fingolimod are not consistent with the progression observed in the trial, with the model predicted progression being faster than that observed in the trial. This is consistent with, and perhaps due to, the differences observed for natural history progression (see Figure 6).

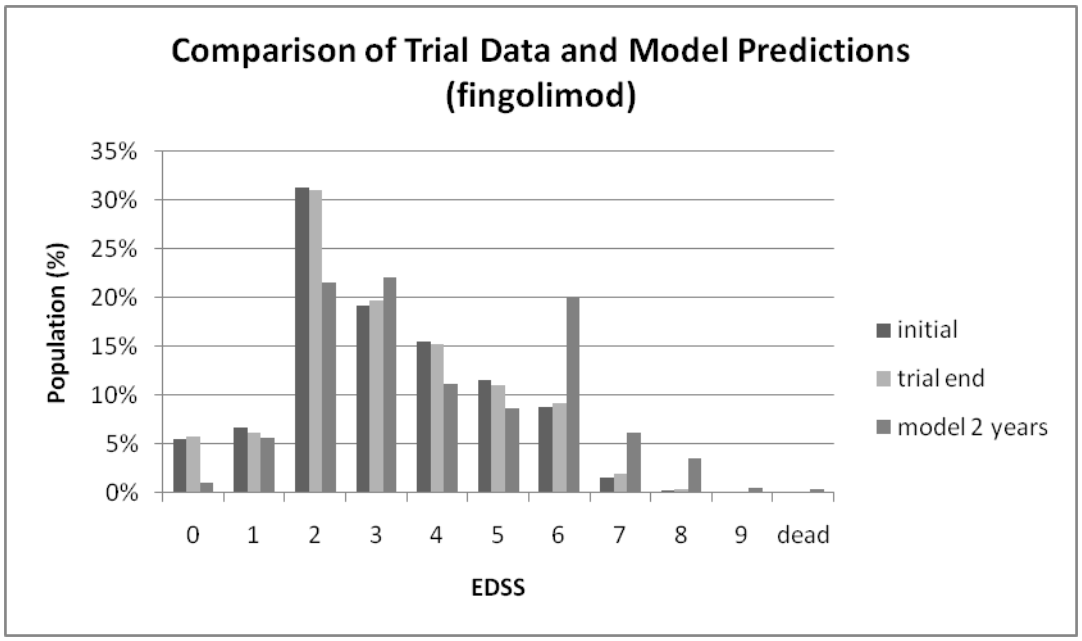

Figure 12: Comparison of Trial Data and Model Predictions for Progression in fingolimod arm (ERG analysis based on manufacturer model)

B - Conversion

The model does not apply a treatment effect to the conversion rate from RRMS to SPMS. Conversion rates are however dependent on EDSS states (Table 15). It is therefore important to note that by modifying progression through these EDSS states, the treatments also indirectly modify the proportions of the cohort converting from RRMS to SPMS over the modelled time horizon. Figure 13 below shows the model predicted percentage of the living population in the SPMS state over time for the two treatments showing a small but noticeable delay in conversion for patients treated with fingolimod as compared to those treated with Avonex.

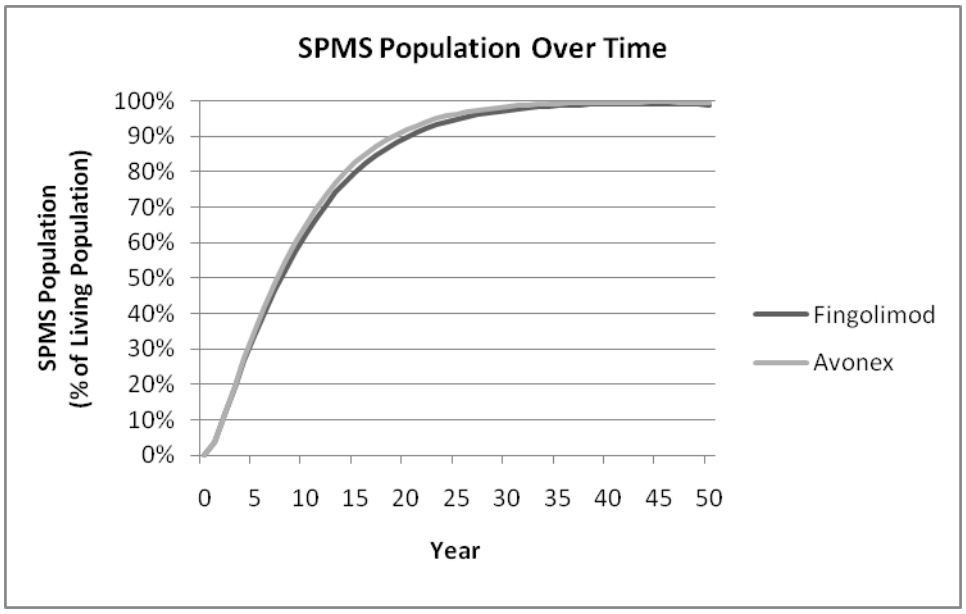

Figure 13: Treatment SPMS Population over Time (ERG analysis based on manufacturer model)

C - Relapses

The effect of treatment on the mean number of relapses is modelled as a relative risk. Table 18 below summarises the relative risk used in the model. A full description of the calculation of these values can be found in Appendix 4. Adjusted annualised relapse rates reported as the primary endpoints in the manufacturer's trials and referred to in the submission were not used in the model.

Table 18: Relative Risk of Relapse

fingolimod vs. placebo	**Avonex vs. placebo**
0.559 (CI: 0.388 to 0.805)	**0.933** (CI: 0.567 to 1.535)

Figure 14 below shows the impact of adjusting the natural history relapse rates in Table 16 with these treatment effects. The graph shows that on average a patient on fingolimod is predicted to experience approximately 12 relapses over the 50 year model time horizon while a patient on Avonex is predicted to experience approximately 14 relapses.

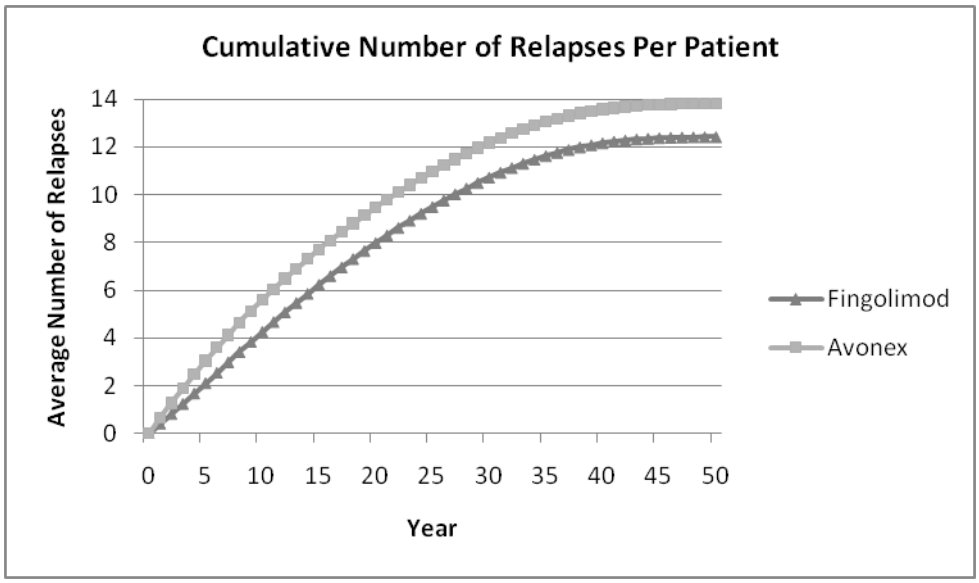

Figure 14: Treatment relapse (ERG analysis based on manufacturer model)

Any treatment that impacts on progression will also have an indirect impact on relapse. This is in addition to any direct impact the treatment has on relapse effect. The correlation between the direct and indirect treatment effects of relapse has not been captured in the model and potentially results in double counting of treatment effect. We will return to exploring the impact of this structural weakness in the manufacturer's model in section 6.

D - Mortality

While the treatments are assumed to have no direct effect on mortality, mortality is partly determined by EDSS state and so by modifying the rate of disease progression the treatments indirectly impact mortality rates. The very marginal difference in mortality for the two treatments as predicted by the model is depicted in Figure 15 below, showing that on average patients on fingolimod are predicted to die very slightly later than those on Avonex with mortality levels converging at 50 years when practically the entire population has died.

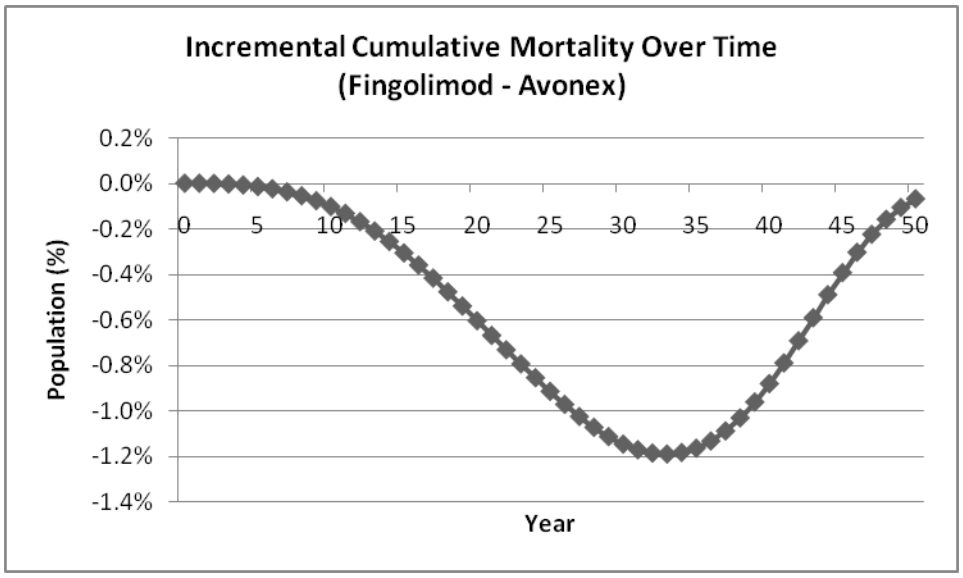

Figure 15: Incremental Mortality (ERG analysis based on manufacturer model)

E - Treatment discontinuation & Adverse Events

One of the reasons for treatment discontinuation is due to AEs. The discontinuations due to AEs are applied in the model throughout the on-treatment period. The data on discontinuations due to AEs are obtained from the head-to-head trial; however, the discontinuation data for the whole trial population are applied rather than for the subset of interest (population 1b). The annual probabilities for discontinuation of treatment used in the model for fingolimod and Avonex are 0.031 and 0.016 respectively. The ERG is of the view that factors such as loss of response and compliance may influence discontinuation rates more in the longer term and may not be captured by these short time horizon trials.

General issues concerning treatment effectiveness

The ERG has a number of significant concerns about the approach used to estimate the relative effectiveness data. In particular, the ERG considers that the current approach excludes potentially relevant trial evidence both for fingolimod and Avonex and, perhaps more importantly, excludes other relevant comparators that were included within the MTC analysis reported in the clinical effectiveness section. The ERG does not consider that the manufacturer provided sufficient justification for ignoring the wider evidence base considered within the MTC and for not exploring more fully the sensitivity of the base-case results to alternative scenarios using relative risk estimates based on alternative approaches and data sources.

Effectiveness estimates for Avonex are derived from an indirect comparison using the FREEDOMS and TRANSFORMS studies; however, the manufacturer's submission refers to the MSCRG trial directly comparing Avonex to placebo which has not been used to inform the effectiveness estimates. There are also a number of other studies referenced that directly compare fingolimod or Avonex to other comparators – these could also be informative as part of a network.

Looking to the manufacturer's MTC, while the estimates from this analysis are prone to criticisms (see Section 1.4), the treatment effects for progression and relapse estimated from this MTC (reproduced in Table 19 below for comparisons of each of the treatments with placebo) give us a sense for the relative efficacies of the different DMTs in a broader MS population. The analyses show that other interferons and glatiramer acetate may be more effective than Avonex. We will explore the impact on cost-effectiveness of using alternative comparators in section 6.

Table 19: MTC Results - Relative risks compared to placebo (Manufacturer's Submission, Section 5.7.6, Table 34 and 35)

Relative risk of treatment	Relative Risk of Progression (95% confidence interval)	Relative Risk of Relapse (95% confidence interval)
fingolimod 0.5 mg	[redacted]	[redacted]
Interferon-beta-1a 22 mcg (Rebif-22)	[redacted]	[redacted]
Interferon-beta-1a 44 mcg (Rebif-44)	[redacted]	[redacted]
Interferon-beta-1a 30 mcg (**Avonex**)	[redacted]	[redacted]
Interferon-beta-1b 250 mcg	[redacted]	[redacted]
Glatiramer acetate 20 mg	[redacted]	[redacted]
Natalizumab 300 mg	[redacted]	[redacted]

The second issue of significant concern to the ERG is the lack of any attempt at validation of the model results. The model predictions have neither been compared to the trial observations nor to any external results. The ERG has attempted to do some independent validation of the model against trial data with the limited data that was made available. The results of this analysis suggest that the model is unable to reproduce outcomes observed in the trials.

The third issue of concern to the ERG is the use of relative risks rather than hazard ratios when adjusting natural history estimates of progression to account for the use of treatments. Adjusting a set of transition probabilities by applying relative risks is incorrect and does not guarantee that the resulting adjusted probabilities are bounded between zero and one. We

will return to this issue in our discussion of the sensitivity analysis and in our additional analyses in Section 6.

The fourth issue of concern to the ERG is around the extrapolation assumptions used in the base case of the model: constant treatment effects derived from the combination of a one year and two year trial are applied over the fifty year time horizon of the model. The ERG feels that this is an important and unjustified assumption and diverges significantly from the previous NICE technology assessments for MS where 20 year time horizons were used. The time horizon was in-fact partially explored in the manufacturer's sensitivity analysis where other more conservative extrapolation scenarios, adjustments to the model time horizon and treatment effect waning, were shown to significantly decrease the treatment effectiveness estimates. Despite these results there was no attempt to justify the structural assumptions used in the base case.

Finally the use of the whole trial population for some model inputs e.g. discontinuation data, whilst using more severe subsets of the trial population for other model inputs, e.g. treatment effects, is not discussed or justified.

Overall, the selective use of data, the lack of validity assessment of results, the unjustified treatment effect extrapolation assumptions and the incorrect usage of relative risks in place of hazard ratios together indicate a high degree of uncertainty around model predictions. Additionally, exploring the wider network of evidence suggests that there may be other more appropriate comparators than Avonex that should have been considered by the manufacturer.

5.2.7 Health Related Quality of Life

The cost-effectiveness model assessed the health-related quality of life (HRQoL) impact of the different treatments for both the patient and their caregivers, and incorporated these in the assessment of quality-adjusted life years (QALYs). In terms of the patients, estimates of HRQoL were used to account for differences in disease progression (EDSS), conversion to SPMS, relapses, treatment disutility and the potential adverse events associated with treatments. The impact on the caregiver was dependent on the EDSS state of the patient.

Health state utilities were applied for each EDSS state in the model. The difference in the cohort over time based on health state occupancy under the treatment alternatives was used to capture the difference in patient utilities. UK EQ-5D based utility weights by EDSS score were derived from Orme et al. (2007), which used data from the 2005 UK MS Survey. The study presents data for RRMS patients.[45]

While EQ-5D data on patient utility was collected as part of the trials, this was not used within the economic model. Instead, the manufacturer used external literature to estimate the relationship between EDSS scores and EQ-5D. No justification was given for choosing external literature in favour of the trial data and, while several external studies were identified, the choice of the Orme et al (2007) study from those identified was not justified. Previous use of the utility data from this study has been heavily criticised in the PenTAG ERG report on natalizumab (TA127), highlighting the low response rates, selection bias, unrepresentative population and self reported severity estimates.[5, 45]

The utility values used in the model are shown in Figure 16. It is important to note the uncertainty around these values, where confidence intervals overlap across almost the entire EDSS state range. The relationship between EDSS score and utility applied in the model is non-linear. For example, the decrease in utility of moving from EDSS 3 to EDSS 7 (-0.277) is smaller than that for moving from EDSS 7 to EDSS 8 (-0.346), additionally the point estimates for utility increase as the disease progresses from EDSS 3 (0.574) to EDSS 4 (0.610). It is important to note that the utilities used in the model include negative utilities for higher EDSS states implying that these are states are worse than death.

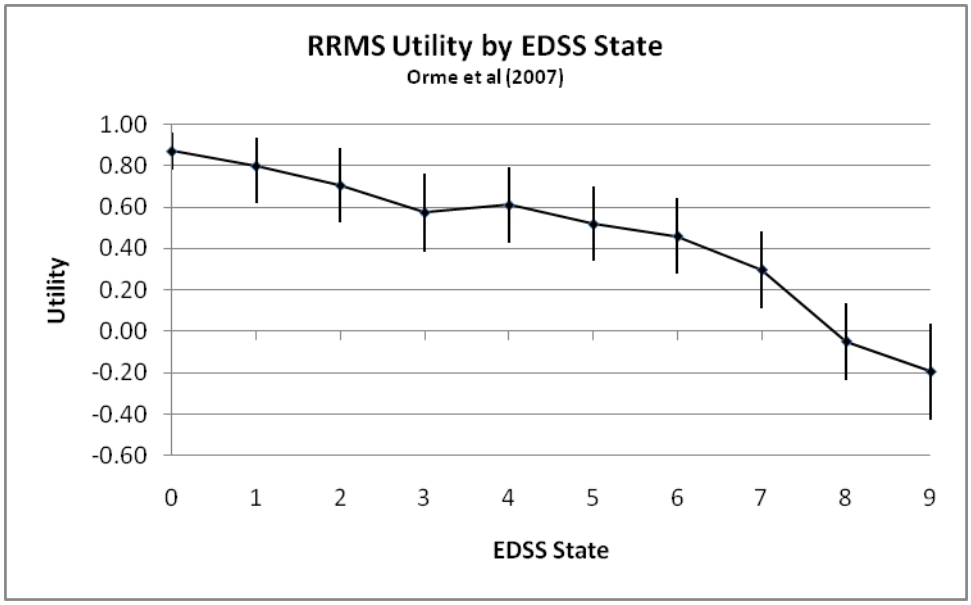

Figure 16: RRMS Utilities by EDSS State (used in model)

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. Adjustments were made to include a utility increment of 0.002 (95% CI: 0.001-0.003; P < 0.001) per year since diagnosis of MS, male gender was associated with an additional utility increment of 0.017 (95% CI: −0.007 to 0.041; P value not stated). A disutility

for recent relapse of −0.071 (95% CI: −0.096 to −0.046; P < 0.001) was also applied. This relapse disutility was applied for 1 year in the model. Disutility of treatment of -0.0345 per year was applied to patients receiving Avonex (Prosser et al. 2003) while the manufacturer assumed that there was no treatment disutility for patients receiving fingolimod (due to this being an oral treatment).[38] Disutility values due to adverse events for fingolimod were applied for macular oedema (annual QALY loss 0.01, prevalence 0.2%) and for atrio-ventricular block second degree (annual QALY loss 0.001, prevalence 0.1%). No disutility values due to adverse events for Avonex are applied in the model.

An estimate of the impact of treatment on the utility of caregivers by patient EDSS state is also used in the analysis. The manufacturer references a previous NICE STA submission (Gani et al 2007) for these values which we have reproduced in Table 20 below.[47]

Table 20: Caregiver Disutility by EDSS State (Manufacturer's Submission, Section 6.4.9, Table 61)

EDSS state	Average hrs of care per patient per day	Average % of day that friends/family spend caring	Weighting relative to maximum disutility*	Disutility of caregivers per patient
0	0.0	0%	0%	0.00
1	0.1	1%	1%	0.00
2	0.3	1%	2%	0.00
3	1.0	45	7%	0.01
4	1.0	4%	6%	0.01
5	2.1	9%	14%	0.02
6	2.9	12%	19%	0.03
7	5.6	23%	38%	0.05
8	11.3	47%	76%	0.11
9	14.8	61%	100%	0.14

There are a number of significant concerns the ERG has with the calculation of the HRQoL data. The first major concern is the fact that while EQ-5D data is collected from the trials for the patient population of interest these data were neither used in the model nor used to validate the utility values that are used in the model. Additionally other external sources of utility data were identified by the manufacturer but subsequently ignored – the selection of the evidence used to inform the model seems arbitrary. The ERG considers that since the submission targets a very specific patient subgroup, it would have been appropriate to use HRQoL data for this same subgroup available directly from the trials.

Table 21 below shows the summary EQ-5D data from the FREEDOMS study. The table suggests that the patients in the placebo arm of the trial experience a [redacted]

whilst those in the fingolimod arm [redacted] experiencing a [redacted] in utility. The results are [redacted]. In contrast to these results, the model predicts for the first two years (to match the two year FREEDOMS trial time horizon) that patients in the placebo arm experience a utility [redacted] and that those in the fingolimod arm do [redacted]. These results lead us to question the face validity of the model predictions, particularly as no justification for this divergence between model predictions and trial observations was provided by the manufacturer.

Table 21: Change from baseline in the EQ-5D in Study D2301 - the FREEDOMS study (Manufacturer's Submission, Section 5.5.3, Table 39)

Change in score from baseline, mean ± SD*	n	Placebo	fingolimod 0.5 mg
EQ-5D utility score	[redacted]	[redacted]	[redacted]

Table 22 and Figure 17 show the utilities by EDSS state derived from the FREEDOMS and TRANSFORMS trials and from the other studies referenced by the manufacturer. The utility data from the trial suggests a flatter relationship, with less disutility associated with progression through EDSS states, as compared to Orme et al (2007).[45] The implication of this is that the effect of DMTs on reducing progression will have a greater impact in QALYs when using the manufacturer selected utility values as compared to utility data observed in the trials.

Figure 17: RRMS Utilities by EDSS State **(Academic in confidence)**

Table 22: HRQoL by EDSS Data from UK EQ-5D Studies (Manufacturer's Submission, Section 6.4.6, Table 59)

EDSS state	Utility estimate								
	Parkin et al., 1998*	Orme et al., 2007*		Biogen Idec UK and Elan Pharma International, 2007		Evidence Review Group Report TA No. 127		FREEDOMS	TRANSFORMS
		RRMS	SPMS	RRMS	SPMS	RRMS	SPMS	RRMS	RRMS
0	—	0.870	0.825	0.91	0.87	0.959	0.874	[redacted]	[redacted]
1	—	0.799	0.754	0.84 (EDSS 0.5-1)	0.8 (EDSS 0.5-1)	0.688 (EDSS 1)	0.603 (EDSS 1)	[redacted]	[redacted]
2	—	0.705	0.660	0.74 (EDSS 1.5-2)	0.7 (EDSS 1.5-2)	0.688 (EDSS 1.5-2)	0.603 (EDSS 1.5-2)	[redacted]	[redacted]
3	0.71	0.574	0.529	0.61 (EDSS 2.5-3)	0.57 (EDSS 2.5-3)	0.645 (EDSS 2.5-3)	0.560 (EDSS 2.5-3)	[redacted]	[redacted]
4	0.66	0.610	0.565	0.65 (EDSS 3.5-4)	0.61 (EDSS 3.5-4)	0.610 (EDSS 3.5-4)	0.527 (EDSS 3.5-4)	[redacted]	[redacted]
5	0.52	0.518	0.473	0.56 (EDSS 4.5-5)	0.51 (EDSS 4.5-5)	0.581 (EDSS 4.5-5)	0.496 (EDSS 4.5-5)	[redacted]	[redacted]
6	0.49	0.460	0.415	0.49 (EDSS 5.5-6)	0.45 (EDSS 5.5-6)	0.538 (EDSS 5.5-6)	0.453 (EDSS 5.5-6)	[redacted]	[redacted]
7	0.35	0.297	0.252	0.44 (EDSS 6.5-7)	0.39 (EDSS 6.5-7)	0.477-0.343 (EDSS 6.5-7)	0.392-0.258 (EDSS 6.5-7)	-	-
8	—	–0.049	–0.094	–0.01 (EDSS 7.5-8)	–0.05 (EDSS 7.5-8)	0.343-0.232 (EDSS 7.5-8)	0.258-0.147 (EDSS 7.5-8)	-	-
9	—	–0.195	–0.240	–0.15 (EDSS 8.5-9.5)	–0.19 (EDSS 8.5-9.5)	0.232 to –0.135 (EDSS 8.5-9.5)	0.147 to –0.220 (EDSS 8.5-9.5)	-	-

As well as there being a range of published studies providing values for EDSS state based utilities, there are also a range of published studies providing values for the disutility of relapse. The manufacturer's submission does not cite any of these studies, however the review by Naci et al (2010) shows this disutility of relapse to vary widely, ranging between -0.0635 and -0.8.[48]

The ERG also has concerns around some of the assumptions used in the utility calculations. It is unclear whether all potential adverse events for fingolimod have been listed. The low incidence of the adverse events that have been listed implies that they will have a negligible utility impact in the model. The source of the incidence data is unclear from the submission. There seems to be no relation in the model between the adverse event data used for utility purposes, which plays almost no part in the model, and the adverse event data used to predict treatment withdrawal.

Overall the ERG feels that there is considerable uncertainty surrounding the utility values used in the model. Data collected in the trials have been not used in the model and, at face value, do not seem to be consistent with the external data used to inform the model. The ERG feels that assumptions around adverse events and treatment disutilities have not been adequately justified. We will explore the impact of using alternative utility data in the model in Section 6.

5.2.8 Resources and Costs

The cost analysis was conducted from an NHS and PSS perspective and considers only resources relevant to the management of the disease. The manufacturer separates the analysis of costs into three sections: drug acquisition costs, administration and monitoring costs, and disease costs. It appears 2010 was used as the costing year.

Disease Costs

The manufacturer's submission considers the cost of each EDSS state, cost of relapse and cost of adverse events. The annual cost to the public health care sector of a patient in each EDSS state is relevant as it accounts for the additional costs incurred due to disability progression. The manufacturer uses data from Biogen Idec UK and Elan Pharma International (2009), as shown in the table below.[49]

Table 23: Disease Cost by EDSS State (Manufacturer's Submission, Section 6.5.6, Table 69)

EDSS State	NHS and PSS annual cost (£, Year 2010)*
0	746
1	1,083
2	1,032
3	3,223
4	2,052
5	2,972
6	3,677
7	8,630
8	20,301
9	19,059

Source: Biogen Idec UK and Elan Pharma International, 2007.
* Inflated to 2010 values using inflation indices derived from Curtis (2010).

No mention is made in the submission that costs associated with patients suffering from RRMS being different to those of SPMS patients. This is in contrast to the TA 127, which assumed a cost per SPMS patient £56 per year higher than RRMS, it is noted that this is a relatively small cost in the scale of general disease costs. The assumption is also made that disease costs are the same regardless of treatment. It is unclear how, or if, the cost per EDSS state data provided in the disease cost section above take account of the cost of adverse events or monitoring costs which may vary between EDSS states.

Alternative sources of cost data by EDSS state were provided in the submission (Manufacturer's Submission, Section 6.5.3, Table 67), these were gathered by a full systematic review of the literature. No attempt is made by the manufacturer to include this in any sensitivity analysis directly. However, the sensitivity of disease by EDSS score is subject to sensitivity analysis through a variation of all of the cost for each EDSS state by a given percentage (Manufacturer's Submission, Section 6.7.7, Table 79). The Natalizumab (TA127) report uses the same source of data for the cost per EDSS state as this submission; however PenTAG (the ERG) raised concerns over the methodology used by the UK MS Survey to gather and extrapolate the data to inform the costs.[5]

The model assumes that, when patients discontinue treatment with DMTs and are provided with BSC, the only costs incurred on the NHS and DSS are the disease costs by EDSS states given in the table above.

The cost of a MS relapse is also included in the manufacturer's submission based on the 2010-2011 National Tariff (Department of Health, 2010) of "admitted patient care &

outpatient procedure tariff, AA30Z multiple sclerosis non-elective tariff", at a cost of £3,039.[50] This is significantly different to the cost per relapse provided in the UK MS Survey (2005), as used in the natalizumab submission (TA127).[5, 45] In that submission the cost per relapse was given as £228 from the NHS and PSS perspective. While cost of relapse is varied in the sensitivity analysis, it is only adjusted to £2,431 as a lower bound, so the lack of impact observed could well be driven by the very limited range of variation allowed. The ERG investigated the valuation of the cost of a relapse and the definitions from the 2010-2011 National Tariff used by the manufacturer to cost relapse.[39] An alternative valuation can be taken from Tyas et al. (2007) which reports the (2005) cost of a relapse as £1,623 (95% CI £554-£2,692).[46] This cost is derived by the authors from previous cost data using seemingly unrelated regression. The ERG would like to highlight that a non systematic approach seems to have been used by the manufacturer to identify relevant cost data, and that the cost associated with relapse may have been overestimated.

Adverse event costs

The cost of treatment related adverse events incurred by patients is provided in the manufacturer's submission for serious adverse events associated with fingolimod (reproduced in the table below).

Table 24: Serious fingolimod Related Adverse Events (Manufacturer's Submission, Section 6.5.7, Table 70)

Adverse events	Items	Unit cost
Macular oedema	Visit to ophthalmologist (1 first attendance)	£105
	Visit to ophthalmologist (1 follow-up attendance)	£74
Atrio-ventricular block, first degree	Non-elective inpatient stay*	£427
Atrio-ventricular block, second degree	Non-elective inpatient stay*	£427

* Based on arrhythmia or conduction disorders with complications.

The adverse event costs associated with Avonex are not included in the base-case analysis the manufacturer states in the submission that they are considered in the sensitivity analysis section, however this does not appear to be the case.

The manufacturer provides a list of serious adverse events that affect patients on fingolimod:

i. Severe infections
ii. Macular oedema
iii. Atrio-ventricular block, first degree
iv. Atrio-ventricular block, second degree

Only costs for adverse events ii, iii and iv are considered in the model. It is unclear why severe infections are excluded from the analysis. Furthermore, it is not clear what adverse event costs are considered for other treatments, or the cost implications of non-serious adverse events on the treatment population. The manufacturer noted that they are still awaiting clarification from the EMA about therapies that may be prescribed to reduce AEs. It is unclear from the submission as to how these additional therapies will affect the analysis, or the scale of additional therapies that may be necessary.

Intervention and comparator drug acquisition costs

Table 25 summarises the drug acquisition costs for fingolimod, Avonex and all other DMTs. Avonex (and many of the other DMTs) is currently made available via the MS Risk Sharing Scheme (RSS). The acquisition costs for Avonex are therefore presented separately based on the current published cost taken from the British National Formulary (BNF 60) and also based on the cost specified as part of the MS RSS.[40] The acquisition costs of the drugs are presented over a year to account for differences in the frequency of drug provision. The mean costs per year for fingolimod and Avonex are broken down in the following sections.

Table 25: Drug Costs used in the Budget Impact Calculations (Manufacturer's Submission, Section 6.5.5, Table 68)

Items	Fingolimod (Gilenya)	Interferon-beta-1a (Rebif) 22	Interferon-beta-1a (Rebif) 44	Interferon-beta-1a (Avonex)	Interferon-beta-1b (Betaferon)	Glatiramer acetate (Copaxone)	Natalizumab (Tysabri)
Mean cost per year (without risk-share)	£19,175	£8,161	£10,623	£8,531	£7,265	£6,841	£14,740
Mean cost per year (risk-sharing)	—	£7,513	£8,942	£8,502	£7,279	£5,823	—

Source: British National Formulary and the UK MS Risk Sharing Scheme (RSS) HSC 2002

Table 25 above highlights the variation in the existing costs of potential comparators to fingolimod, even under the risk-sharing scheme. This is important in highlighting the impact of the manufacturer only considering Avonex as a comparator as in many cases the alternative would be a cheaper DMT.

The ERG would like to raise their concern over potential inaccuracies in the submission's administration costs associated with natalizumab. The submission reports administration costs associated with natalizumab of £16,861, this is more than twice the administration cost provided in the NICE costing template for natalizumab (of £8,379). While natalizumab is not used as a comparator to fingolimod, the ERG believes that such inaccuracies limit the potential for an accurate consideration of all of the possible alternatives. The justification given by the manufacturer in their response to the points for clarification is that the HRG

code used for the original costing template (A18) has been superseded by AA30Z; it is unclear to the ERG if this analysis is correct or explains the entire difference in cost.

Administration and monitoring costs

Due to the fact that that fingolimod is administered orally it is not subject to the administration costs of the comparators, which are all injectable treatments. The administration costs of Avonex are assumed to be a one off cost of £78 to train the patient to self-administer.

The manufacturer provided drug costs of fingolimod including the administration and monitoring costs of the treatment, as well as tests necessary to assess patient suitability for treatment with fingolimod. Table 26 below highlights the additional administration and monitoring costs associated with fingolimod, these are:

- A six hour observation of all patients during first administration.
- Additional ophthalmologist visits (one for all patients at treatment initiation then one a year follow up for 0.9% of patients).
- Additional tests including pregnancy, basic metabolism and chicken pox pre-exposure.

The requirement for these additional tests is based on the SPC for fingolimod. The frequency of patients who need any additional resources associated with the SPC requirements is taken from FREEDOMS.

Table 26: Administration and Monitoring Costs of fingolimod and Avonex (Manufacturer's Submission, Section 6.5.5, Table 68)

Drug/costs related to administration	Unit cost (£)	Requiring	Annual resource use (units)	
			First year	Sub-sequent year
Fingolimod				
Physician visit				
-Neurology visit	206.53	100%	2	1
-Ophthalmology visit (treatment initiation)	105.47	100%	1	0
-Ophthalmology visit (follow-up)	73.84	0.9%	1	0
Tests/imaging				
-Full blood count	3.06	100%	4	2
-Liver function	1.29	100%	4	2
-Pregnancy test	1.29	69%	1	0
-Basic metabolism	1.29	100%	2	2
-Test for prior exposure to chicken pox	7.25	10%	1	0
Other				
-Patient observation following first administration	501.43	100%	1	0
-Protocol-mandated hospitalization	2,078.68	2%	1	0
-AV block requiring atropine	0.68	0.2%	1	0
-Evaluation of the fundus	105.47	3.5%	1	0
Avonex				
Injection administration visits				

-Self-administration (training)	78.00	100%	1	0
Physician visits				
-Neurology visit	206.53	100%	4	1
Tests/imaging				
-Full blood count	3.06	100%	4	2
-Liver function	1.29	100%	4	2

Source: SPC and guidelines from Association of British Neurologist 2009 MS guidelines

While the administrative and monitoring costs of fingolimod and Avonex are fully provided by the manufacturer (as highlighted in the above table), they are not well justified. A specific example of the lack of clarity in the costs is the provision of 4 neurology visits to Avonex patients in the first year of treatment but only 2 for Fingolimod. It is not clear to the ERG why the number of visits varies between treatments.

5.2.9 Cost-Effectiveness Results

The base case cost-effectiveness results presented in the manufacturer's submission, and reproduced below in Table 27, are based on the deterministic estimates from the model. The results are produced for the pooled non-responder patient population (population 1b) from the FREEDOMS and TRANSFORMS trials. The results are shown for a 50 year model time horizon with a discount rate of 3.5% applied to both costs and effects. The ICER from this deterministic model for fingolimod compared to Avonex is £55,634 per QALY.

Table 27: Discounted Deterministic Cost-Effectiveness Results (Manufacturer's Submission, Section 6.7.6, Table 78)

Tech-nologies	Total costs (£)	Total QALYs	Incre-mental costs (£)	Incre-mental QALYs	ICER £ per QALY gained
Avonex	271,647	3.98	—	—	—
Fingolimod	321,721	4.88	50,084	0.90	55,634

ICER, incremental cost-effectiveness ratio; LYG, life-year gained; N/A, not available; QALY, quality-adjusted life-years.

Figure 18 below produced by the ERG shows how the ICER evolves over the model time horizon. We can see from the figure that the ICER decreases over time though never falls under commonly used threshold values.

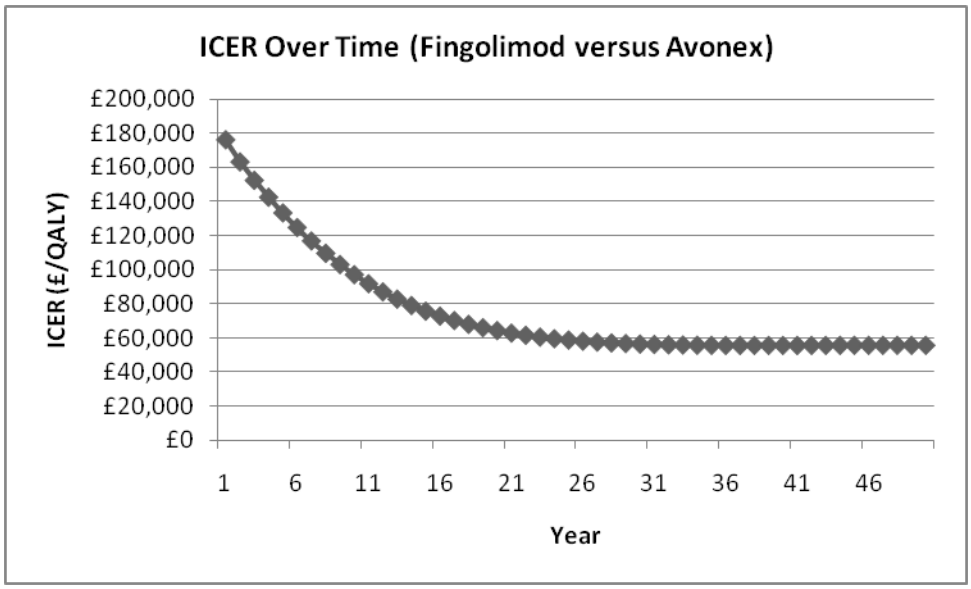

Figure 18: fingolimod versus Avonex ICER over time (ERG analysis based on manufacturer base case model, evaluated deterministically)

The manufacturer has built a probabilistic model, but cost-effectiveness results presented in the submission are deterministic (obtained by evaluating the model using mean parameter values). Whilst deterministic results would provide a suitable approximation for a linear model, for this complex non-linear model it is necessary to instead use probabilistic results. Table 28 below shows probabilistic results from the model calculated using the manufacturer's suggested parameter distributions and averaged over 5,000 model iterations.

Table 28: Discounted Probabilistic Cost-Effectiveness Results (ERG analysis based on manufacturer model)

Tech-nologies	Total costs (£)	Total QALYs	Incre-mental costs (£)	Incre-mental QALYs	ICER (£) (QALYs)
Avonex	271,469	3.89	—	—	—
Fingolimod	322,562	4.63	51,093	0.73	69,787

ICER, incremental cost-effectiveness ratio; LYG, life-year gained; N/A, not available; QALY, quality-adjusted life-years.

The probabilistic analysis results in a higher ICER of £69,787 per QALY (compared to the deterministic estimate of £55,634 per QALY).

5.2.10 Sensitivity Analyses

Sensitivity analysis has been conducted in the model to represent parameter uncertainty and uncertainty for key structural assumptions. Both deterministic sensitivity analysis (where one parameter is varied at a time) and probabilistic sensitivity analysis (PSA - where all parameters are varied simultaneously) were conducted. The results of the deterministic sensitivity analysis are shown in Table 29 below. There is no discussion of or justification for

the parameters varied in the sensitivity analysis or the range over which these parameters are varied. It is clear from the table that the relative risks of progression are by far the most significant sources of uncertainty (of those that were explored). The ICER in the table ranges from as low as £6,132 per QALY to fingolimod being dominated (i.e. less effective and more costly), demonstrating the large degree of uncertainty in the model.

Table 29: Deterministic Sensitivity Analysis (Manufacturer's Submission, Section 6.7.7, Table 79)

Parameter		Level	Value	ICER
Efficacy	RR of progression for fingolimod	Lower 95% CI	0.332	£24,686
		Upper 95% CI	1.210	-£107,276
	RR of progression for Avonex	Lower 95% CI	0.308	-£75,683
		Upper 95% CI	2.404	£6,132
	RR of relapse for fingolimod	Lower 95% CI	0.388	£50,500
		Upper 95% CI	0.805	£64,107
	RR of relapse for Avonex	Lower 95% CI	0.567	£68,880
		Upper 95% CI	1.535	£39,558
	Discontinuation rate for fingolimod	Lower 95% CI	0.0045	£61,265
		Upper 95% CI	0.0342	£55,030
	Discontinuation rate for Avonex	Lower 95% CI	0.0138	£55,074
		Upper 95% CI	0.0545	£62,312
Cost	Cost of relapse	80% of base values	£2,431	£56,495
		120% of base values	£3,647	£54,773
	Cost of disease by EDSS stage	80% of base values	£597 to £16,241	£57,772
		120% of base values	£895 to £24,361	£53,495
Utility	Utility of EDSS stages	80% of base values	RRMS: 0.696 to –0.125 SPMS: 0.660 to –0.161	£63,990
		120% of base values	RRMS: 1 to –0.188 SPMS: 0.990 to –0.241	£49,279
	Utility adjustment from years since diagnosis	Lower 95% CI	0.001	£55,851
		Upper 95% CI	0.003	£55,418
	Utility adjustment for males	Lower 95% CI	–0.007	£55,682
		Upper 95% CI	0.041	£55,586
	Disutility of relapse	Lower 95% CI	–0.096	£53,731
		Upper 95% CI	–0.046	£57,676
	Disutility of treatment	80% of base values	–0.0079 to –0.0383	£58,418
		120% of base values	–0.01188 to –0.05742	£53,103
	Discounting rate	Lowest value	0%	£43,197
		Highest value	6%	£64,340

CI, confidence interval; EDSS, Expanded Disability Status Scale; RR, relative risk.

The manufacturer also conducted a PSA. How this was done is not described in the submission, though from the ERG's analysis of the model it seems that the same parameters considered in the deterministic sensitivity analysis were varied. The selection of distributions for these parameters was not discussed or justified. Figure 19 below shows incremental costs and effect pairs for each of the 5000 iterations of the PSA from the manufacturer's submission. We can see from the figure that throughout the PSA fingolimod is always more costly than Avonex and more effective than Avonex in approximately ¾ of the iterations.

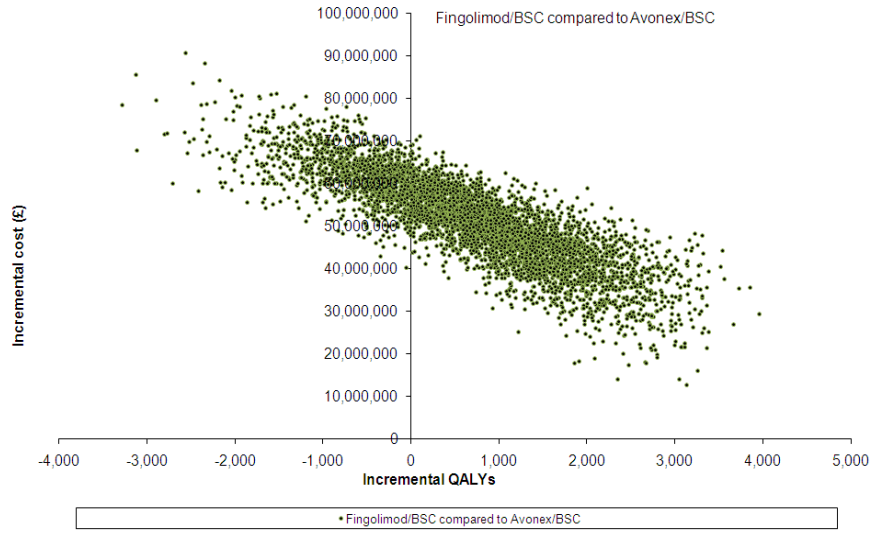

Figure 19: PSA Scatter Plot (Manufacturer's Submission, Section 6.7.8, Figure 19)

The results of the PSA are summarised in the cost-effectiveness acceptability curve (CEAC) in Figure 20 below. From the figure we can see that 26% of iterations from the PSA fell below £30,000 per QALY and 50% of iterations fell under £68,000 per QALY.

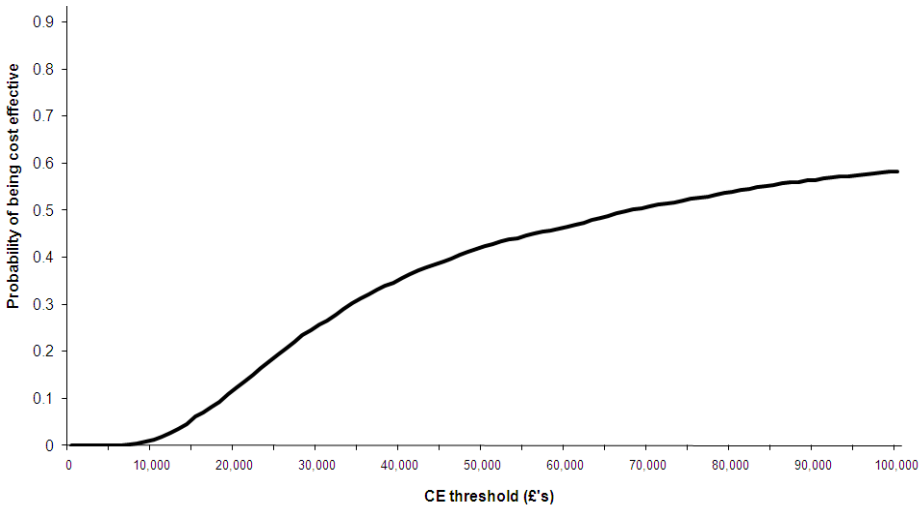

Figure 20: Cost-Effectiveness Acceptability Curve (Manufacturer's Submission, Section 6.7.8, Figure 20)

As mentioned before, the manufacturer used relative risks rather than hazard ratios to account for treatment effects; this creates issues in the implementation of the PSA analysis such that in the course of the PSA the individual adjusted probabilities can assume a value higher than 1. Moreover, the probability of leaving a certain state can be assigned a negative value. The manufacturer attempted to resolve this issue by discarding the iterations of the PSA that returned negative probabilities. By doing so, the resulting PSA does not fully capture the uncertainty in the relative effect parameter estimates – the distributions of these being effectively truncated to accommodate the limitations of using relative risks in place of hazard ratios.

Some structural parameters were also examined in the sensitivity analysis. Of these structural parameters examined only those dealing with treatment effect time horizon and model time horizon had significant impacts on the cost-effectiveness estimates. The manufacturer limited their analysis of these structural scenarios to deterministic results, but the ERG felt it important for this non-linear model to augment these with the corresponding probabilistic results; these are shown in Table 30 below. It is clear from the table that alternative structural assumptions can significantly impact the ICER. While the ERG recognises that previous NICE technology appraisals have used a 20 year time horizon it is felt that the 50 year time horizon chosen better represents the lifetime perspective for the patients in the model as was seen in Figure 10. The ERG, however, does consider waning of treatment effect to be a useful tool in exploring more realistic extrapolation assumptions and further analysis along these lines will be returned to in section 6.

Table 30: Structural Sensitivity Analysis

Structural Parameter / Assumption		ICER (£) Deterministic Model	ICER (£) Probabilistic Model[3]
Waning of treatment effect	Continued treatment effect	£55,634	£69,787
	Waning at 2 years: reduction to 50% efficacy levels	£73,191	£83,851
	Waning at 2 years: reduction to 25% efficacy levels	£85,266	£95, 312
	Waning at 5 years: reduction to 50% efficacy levels	£63,890	£74,923
	Waning at 5 years: reduction to 25% efficacy levels	£68,493	£78,024
Time Horizon	10-year horizon	£97,159	£105,221
	20-year horizon	£64,280	£76,843
	30-year horizon	£56,368	£68,512
	40-year horizon	£55,556	£70,373

There were several other parameters that the ERG would have liked to see explored as part of the sensitivity analysis. In particular, it would have been useful to consider alternative scenarios around the natural history data used, especially as there were many different potential sources of data; the selection among which in the model seemed to a large extent to be arbitrary. Specific scenarios using alternative utility data sources would also have been reassuring to see given the high degree of uncertainty around the utility data used. We will return to some of these scenarios in section 6.

Despite the large degree of uncertainty in the model demonstrated by the sensitivity analysis there is no critical analysis of the model, the parameter values or the structural assumptions in light of these results. This is particularly concerning where structural assumptions and parameter values used differ as compared to those used in previously published NICE technology appraisals for MS, with no justification given for this divergence.

5.2.11 Model Validation

Model validation was dealt with by the manufacturer in a limited manner and seems to have been restricted to checking that formulas in spreadsheets were typed correctly and ensuring referenced values were copied into the model correctly. Some extreme value testing of the model was carried out to ensure that the model behaved as it was described to in the submission e.g. setting mortality rate to 0 and observing that the model predicts no deaths.

No internal validation or goodness of fit was reported comparing the model predictions with observations from the FREEDOMS, TRANSFORMS or Ontario studies, the three main sources of evidence used to inform the model. No external validation was reported

[3] probabilistic estimates based on ERG analysis of manufacturer model

comparing the model predicted results with other trials or other published model results. Clinical expertise was not utilised to verify that the model captured a plausible abstraction of the disease or made clinically plausible predictions.

The ERGs have attempted to check the model predictions against the results observed in the trials throughout the current section and have consistently found divergences between them.

5.2.12 Additional Subgroup Analysis (1b but not 2)

The ERG considers the choice of the 1b population as being problematic, as it contains a mixture of RES and non-RES patients. These two subpopulations have different treatment options available to them and hence should be treated separately (as separate decision problems). In response to this concern the manufacturer provided subgroup data for the non-RES subset of the 1b population (1b but not 2 population). It is interesting to note that the overlap between this new non-RES population and the original 1b trial subpopulations across the 4 arms of the trials in the analysis ranges from [redacted] to [redacted]. Despite this fact, the relative treatment effects estimated using evidence from the trials on both subpopulations differ significantly (Table 31).

Table 31: Treatment Effect Relative Risks non-RES subgroup of 1b Population

	fingolimod vs Placebo	Avonex vs Placebo
RR Progression	[redacted]	[redacted]
RR Relapse	[redacted]	[redacted]

The cost-effectiveness results that the manufacturer has submitted resulting from these changes are reproduced in Table 32.

Table 32: Deterministic Cost-Effectiveness Results non-RES subgroup of 1b Population

Tech-nologies	Total costs (£)	Total QALYs	Incre-mental costs (£)	Incre-mental QALYs	ICER (£) (QALYs)
Avonex	278,328	2.98	—	—	—
Fingolimod	316,748	5.03	38,420	2.05	18,741

ICER, incremental cost-effectiveness ratio; LYG, life-year gained; N/A, not available; QALY, quality-adjusted life-years.

The relative risk of progression for Avonex relative to placebo reported in Table 31 above suggests that Avonex is significantly worse than even placebo in this population, supporting our concerns (highlighted in earlier sections) around the suitability of Avonex as the sole

comparator. It is interesting to note that the level of uncertainty around this estimate has also increased significantly. The results reproduced in Table 32 should be interpreted with this in mind. More appropriate comparisons in this non-RES population will be discussed in Section 6.

5.2.13 Key Critiques and Uncertainties

The ERG has a number of general concerns about the approach taken in the submission:

- The manufacturer does not appear to have used a systematic approach to identify and subsequently select appropriate data sources to inform the key parameters of the model – choices of data appear to be arbitrary and unjustified.
- Methods used for deriving the various model parameters from the selected data are not fully described and assumptions made in using these methods are not discussed or justified.
- There has been no attempt by the manufacturer to validate the predictions of the model either internally against the trial data or externally against other published studies or clinician opinion.

More specific concerns arising from the various sections of the analysis are highlighted in Table 33 below.

Table 33: Key Critiques and Uncertainties

Section	Critique	Impact on Cost-Effectiveness	Explored in Section 6
5.2.2 Population	The manufacturer's submission only provides cost-effectiveness results for the sub-population 1b (as defined by the manufacturer). The significant level of cross-over between the populations considered in groups 1a, 1b and 2 is not dealt with or discussed.	High	Yes
	The manufacturer states that the 'model's population [represented by the pooled analysis of FREEDOMS and TRANSFORMS] appears to be reflective of the population eligible to receive disease-modifying therapies in England and Wales'. The distribution of patients across EDSS states however, seems implausibly low with a significant proportion of the population having an EDSS score of 0.	High	Yes
5.2.3 Interventions	The validity of the use of Avonex to represent all interferon-beta drugs is not well justified, nor is a	High	Yes

Section	Critique	Impact on Cost-Effectiveness	Explored in Section 6
and Comparators	sensitivity analysis run to consider the implications of using an alternative comparator.		
	Natalizumab is excluded from the analysis, despite its use in patients who have RES RRMS.	Low	No
	The manufacturer makes the assumption that treatment efficacy is equivalent for BSC as for placebo, this is not justified or the implications of it tested.	Unknown	No
5.2.4 Model Structure	Concerns have been raised in the literature as to the validity of the EDSS method to characterise different levels of Multiple Sclerosis severity.	Unknown	No
	The assumption of the same mortality rate applying to RRMS and SPMS is not well justified.	Low	No
	Correlation between relapses and disability progression was not captured in the model.	Medium	Yes
5.2.5 Natural History	The model uses and adjusted version of the Ontario dataset, the adjustments made are not clearly described. This makes the results difficult to generalise and compare with previous NICE technology appraisals.	Medium	Yes
	The lack of correlation between progression and relapse means that any relative risk applied to these will double count benefits.	Medium	Yes
	Trial data largely ignored - used neither in the model natural history nor to validate model natural history inputs.	Unknown	No
	Model predictions inconsistent with trial data as analysed by the ERG.	Unknown	No
	Data sources used to inform natural history parameters selected seemingly arbitrarily, no systematic process followed.	Unknown	No

Section	Critique	Impact on Cost-Effectiveness	Explored in Section 6
	Calculations used to derive natural history parameters are not transparent and assumptions used in these are not justified.	Unknown	No
5.2.6 Treatment Effectiveness	Much of the available evidence that could have been used to inform effectiveness estimates is ignored. Only the FREEDOMS and TRANSFORMS trials are included.	Medium	Yes
	MTC suggests that Avonex is not the best comparator with other DMTs appearing to be relatively more effective.	High	Yes
	Relative risks are use in place of hazard ratios in the model, these are inappropriate and can result in illogical predictions such as negative transition probabilities.	Medium	Yes
	The use of an initial model patient population distributed disproportionately across lower EDSS states seems to account for some of the progression related benefit of fingolimod.	Medium	Yes
	Extrapolation of a constant treatment effect over 50 years is questionable, especially in light of previous technology appraisals use of a 20 year time horizon for MS. More conservative extrapolation scenarios should be explored.	Medium	Yes
	No validation of model predictions against trial results. ERG's attempt to do this suggest the two do not match.	Unknown	No
	Inconsistent use of trial data where subsets are selectively used.	Low	No
5.2.7 Health Related Quality of Life	HRQL Data were collected in trial but these trial informed values are not used in the model	High	Yes
	Many alternative external HRQL data sources are available the choice of Orme has not been justified and seems to be the most favourable to fingolimod.	High	Yes
	Utility data used in the model is very uncertain and at times suggests counter intuitive results e.g. progressing to higher EDSS states increases utility.	Unknown	No
5.2.8 Resources and Costs	Severe infections and infestations listed in table 42 of the submission are excluded from the model with no justification.	Low	No

Section	Critique	Impact on Cost-Effectiveness	Explored in Section 6
	It is unclear how BSC is costed relative to DMTs.	Unknown	No
5.2.9 Cost-Effectiveness Results	Deterministic results presented, however for a non-linear model such as this one probabilistic results are more appropriate.	High	Yes
	Avonex is not cost-effective in this patient subgroup, BSC or other alternative comparators should be considered in the analysis.	High	Yes
5.2.10 Sensitivity Analyses	Parameters included and ranges of values used in the deterministic sensitivity analysis were not justified.	High	Yes
	Probabilistic sensitivity analysis was not adequately described and distribution of parameters was not justified.	Low	No
	Probabilistic results were not used in the calculation of the ICER and the CEAC was not utilised as a means of quantifying the uncertainty around the ICER.	High	Yes
	Many key sources of uncertainty e.g. the natural history parameters were not meaningfully explored as part of the sensitivity analysis.	High	Yes
	The use of relative risks in place of hazard ratios meant that uncertainty around parameter estimates was not fully explored in the sensitivity analysis and resulted in input parameter distributions being truncated.	Medium	Yes
	More conservative structural assumptions around treatment effect extrapolation and model time horizon are shown to have significant impacts on cost-effectiveness results. The structural assumptions used are not justified in light of these results.	High	Yes
5.2.11 Model Validation	No meaningful internal, external or expert validation of the model and its predicted results was conducted.	Unknown	No
	Where the ERG has attempted to compare model predictions with results from the manufacturer's trials these seem to show substantial divergence between the model predictions and the observed values.	Unknown	No

6 ADDITIONAL ERG ANALYSIS

This section details the ERG's further exploration of some of the issues and uncertainties raised in the review and critique of the manufacturer's cost-effectiveness analysis presented in Section 5. In the initial sub-section, corrections are applied to the manufacturer's model where methodological errors have been found. Further sub-sections explore the various uncertainties highlighted in Section 5 in light of results from the corrected base case developed in the initial sub-section. The structure of this section broadly follows that of Section 5:

- Section 6.1 - Corrected Base Case: Probabilistic analysis using hazard ratios is implemented in place of deterministic analysis using relative risk.
- Section 6.2 - Population: The impact of using alternative EDSS distributions for the initial population is explored.
- Section 6.3 - Interventions and Comparators: Rebif-44 is added as a comparator in the analysis using both data from the head to head EVIDENCE trial and data from the manufacturer's MTC.
- Section 6.4 - Model Structure: Correction for the double counting of treatment effect on relapse is explored by turning off direct treatment effect on relapse and observing only the indirect treatment impact.
- Section 6.5 - Natural History: The effect of varying the natural history progression rates used in the model are explored.
- Section 6.6 - Treatment Effectiveness: Alternative extrapolation scenarios regarding the extrapolation of treatment effect are considered.
- Section 6.7 - Health Related Quality of Life: The use of trial based patient utility data in the model in place of external data sources is explored.
- Section 6.8 - Summary

The ERG expressed concerns about the level of heterogeneity in the target population in the submission (referred to by the manufacturer as population 1b). This population consists of patients not responding to a beta-interferon and includes both rapidly evolving severe (RES) and non-RES patients. The manufacturer also details a RES population (referred to by the manufacturer as population 2) that includes but is not limited to non-responders as defined in the 1b population. The ERG considers it valuable to separate out these two partially overlapping populations as there are different treatment recommendations for the RES and non-RES subgroups; hence different comparators should be included in the cost-effectiveness analysis of the two distinct subgroups. The manufacturer, upon request from the ERG, submitted additional data for the non-RES subset of the non-responder population

(referred to by the manufacturer as population 1b but not 2). In the remainder of the analysis, where possible, results will be presented for both the original 1b population dealt with in the manufacturer's submission and the 1b but not 2 population provided as part of the clarifications from the manufacturer.

The ERG feels it is important to highlight the central role of treatment under best supportive care (BSC) in the model. All comparisons between Avonex and fingolimod in the model are derived indirectly through comparison of the relative effect of each treatment against placebo, as a proxy for BSC. Despite treatment under BSC being so integral to the results produced from the model, BSC is not explicitly included as a comparator in the cost-effectiveness analysis. In addition to this the cost-effectiveness of Avonex in the patient subgroups of interest has not been previously explored. Avonex, being one of the first-line treatment options that these patients failed to respond to, may not be cost-effective as a second-line treatment for these patients. Given this possibility, comparisons solely against Avonex may not be appropriate. For these reasons fully incremental results including explicit consideration of BSC have been presented throughout this section.

6.1 Corrected Base Case

As was highlighted in Section 5, the ERG has identified inaccuracies in the manufacturer's submitted base case: the use of deterministic results (rather than probabilistic) in calculating cost-effectiveness and the use of the relative risks (rather than the hazard ratios) to describe relative effectiveness. These were corrected by the ERG; results and methods used are reported in this section.

Deterministic results (evaluation of the model at mean values for input parameters) do not appropriately estimate cost-effectiveness when the model is non-linear. The manufacturer's model is non-linear due to its Markov structure, thus results should be derived using probabilistic methods (repeatedly drawing from the input parameter distributions and averaging model results across iterations) to derive theoretically sound estimates. All results presented in this section, where possible, will be calculated using a probabilistic evaluation of the model.

In the submission, the manufacturer presented appropriately adjusted hazard ratios to describe the relative effect on progression of treatment with fingolimod in relation to Avonex. The analysis used to generate these estimates has been published in peer reviewed journals. In the model, however, these estimates are ignored and instead relative risks from unadjusted trial data are used. As well as being based on unadjusted data, the approach used to apply the relative risks means that at times (particularly in the course of the PSA) the

probabilities of progression used in the model can assume values higher than one (or lower than zero). The use of hazard ratios avoids the derivation of probabilistically incoherent transition matrices. The ERG has re-analysed the manufacturer's base case using hazard ratios instead of relative risks; the new results represent a more appropriate application of the treatment effect when applying PSA in the model.

To undertake this analysis, the ERG needed to identify the appropriate hazard ratios for the two treatments and apply these to the natural history evidence to derive progression in the model. The manufacturer's submission reports hazard ratios for fingolimod against Avonex from the TRANSFORMS trial, and for fingolimod against placebo (assumed to represent BSC in the model) from the FREEDOMS trial. The ERG derived the hazard ratio for Avonex against placebo using the adjusted indirect comparison method (Bucher et al., 1997).[51] The probability of progressing (from each EDSS starting state) was calculated from the natural history transition matrix (by summing the probabilities of moving from a particular EDSS state to more severe EDSS states). The relevant hazard ratios were then applied to this set of probabilities of progressing under natural history, to obtain probabilities of progressing for the treatments of interest, fingolimod and Avonex. These adjusted probabilities of progressing were then redistributed across the state transition matrices for the two treatments (split back out to represent progression to specific EDSS states) in the proportions observed under natural history. All results presented in this section, where possible, will be derived using the hazard ratio approach described here.

The hazard ratios for progression for each treatment relative to BSC, used by the ERG in the remainder of this section, are shown in Table 34 below.

Table 34: Hazard Ratios for Progression Relative to BSC

	Population 1b (95% confidence interval)	Population 1b but not 2 (95% confidence interval)
Avonex vs. Placebo	[redacted]	[redacted]
Fingolimod vs. Placebo	[redacted]	[redacted]

The ERG has calculated incremental cost-effectiveness results, including BSC as a comparator, from the probabilistic evaluation of the model, using the hazard ratios quoted in Table 34 above for populations 1b and 1b but not 2. To do this mean lifetime costs and QALYs of all strategies are presented and their cost-effectiveness compared simultaneously by estimating ICERs as appropriate using standard decision rules.[52] The ICER examines the additional costs that one strategy incurs over another and compares this with the additional benefits. When more than two strategies are being compared the ICERs are calculated using the following process:

i. The strategies are ranked in terms of cost (from the least expensive to the most costly).
ii. If a strategy is more expensive and less effective than the previous strategy, then this strategy is said to be dominated and is ruled out and excluded from the calculation of the ICERs.
iii. The ICERs are calculated for each successive alternative, from the cheapest to the most costly. If the ICER for a given strategy is higher than that of the next more costly strategy, then this strategy is ruled out on the basis of extended dominance.
iv. Finally, the ICERs are recalculated excluding any strategies that are ruled out using the notions of dominance and extended dominance.

Base case cost-effectiveness results for population 1b and population 1b but not 2 are shown in the tables below.

Table 35: Base Case Cost-Effectiveness Results - Population 1b

	Total Cost (£)	Total QALYs	ICER (£ per QALY)
BSC	224,192	3.66	-
Avonex	272,454	3.76	ED (ICER of 471,431 vs BSC)
fingolimod	321,995	4.70	**94,094**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 36: Base Case Cost-Effectiveness Results - Population 1b but not 2

	Total Cost (£)	Total QALYs	ICER
BSC	219,865	3.64	-
Avonex	274,611	3.06	D
fingolimod	316,649	4.83	**81,369**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

We can see from the base case results presented above that Avonex is dominated or extendedly dominated in both the populations considered. This indicates that BSC rather than Avonex is the appropriate comparator in the cost-effectiveness analysis. The incremental cost-effectiveness ratio for fingolimod in relation to BSC is higher in population 1b (£94,094 per QALY gained) than it is in population 1b but not 2 (£81,369 per QALY gained). The ERG would like to highlight that these corrected base case results are both significantly higher that the ICERs (relative to Avonex) reported in the manufacturer's submission and clarifications: £65,634 for population 1b and £18,741 for population 1b but not 2.

6.2 Population

The ERG had concerns regarding the impact and representativeness of the initial EDSS score distribution used in the model, as discussed in section 5 of this report. The distribution of patients across EDSS states in the model is taken from the distributions of the relevant subgroups (defined post-hoc) of the trials evaluating fingolimod. There is no assessment of the appropriateness of this distribution or comparison with what is observed in clinical practice reported in the submission. An evaluation of the impact of alternative assumptions on the initial EDSS distribution was not conducted by the manufacturer. The ERG has thus decided to explore the impact of initial EDSS distribution on the model results. A number of scenarios were examined, these entailed: the entire population starting in EDSS 2 (the median EDSS value from the trials for both the subgroups presented), as well as EDSS states 3, 4 and 5. The results of these analyses for population 1b and population 1b but not 2 are presented in Table 37 and Table 38 below.

The results show that the cost-effectiveness of fingolimod varies substantially depending on the initial distribution of patients across EDSS states – the ICER ranges from £78,338 to £102,718 in population 1b and from £60,887 to £93,412 in population 1b but not 2. Since the model is clearly sensitive to the initial population EDSS distribution, the assumptions undertaken in the analyses need to be carefully clinically evaluated for plausibility before meaningful interpretations can be made about the cost-effectiveness of fingolimod.

Table 37: Initial Population Distribution Scenarios - Population 1b

	All Population Starts at EDSS 2			All Population Starts at EDSS 3			All Population Starts at EDSS 4			All Population Starts at EDSS 5		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	209,526	4.31	-	242,330	2.43	-	268,583	0.86	-	288,349	-0.26	-
Avonex	261,344	4.35	ED (ICER of 1,226,717)	284,397	2.51	ED (ICER of 493,819)	300,140	1.01	ED (ICER of 210,458)	314,180	-0.13	ED (ICER of 211,330)
fingolimod	312,846	5.31	**102,718**	328,778	3.32	**97,242**	335,802	1.72	**78,338**	344,566	0.44	**80,779**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 38: Initial Population Distribution Scenarios - Population 1b but not 2

	All Population Starts at EDSS 2			All Population Starts at EDSS 3			All Population Starts at EDSS 4			All Population Starts at EDSS 5		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	206,520	5.73	-	236,641	4.11	-	262,744	2.74	-	281,416	1.51	-
Avonex	266,239	5.04	D	285,301	3.53	D	300,546	2.29	D	312,148	1.15	D
fingolimod	308,670	6.83	**93,412**	322,676	5.13	**83,923**	328,993	3.83	**60,887**	336,797	2.38	**63,439**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

6.3 Interventions and Comparators

Adding BSC explicitly as a comparator in the cost-effectiveness results, as demonstrated in the revised base case above, resulted in Avonex being dominated or extendedly dominated as a treatment option. In addition to this, Avonex also has a low market share compared to other beta-interferons (Manufacturer's Submission, Section 2.6, Table A8) and there is evidence that it may be less effective than other treatment options (Manufacturer's Submission, Section 5.7.6, Table 34 and EVIDENCE).[14]

The ERG identified Rebif-44 as an alternative beta-interferon with both higher market share and potentially greater efficacy than Avonex (EVIDENCE).[14] In this section the ERG will present incremental cost-effectiveness results with Rebif-44 included as an additional comparator in the analysis, using alternative input data: either (i) using head-to-head evidence on the effectiveness of Rebif and Avonex, or (ii) using the results of the MTC provided by the manufacturer. The results from these analyses should be interpreted in light of the usual caveats regarding indirect comparisons.

(i) adding Rebif as a comparator using evidence from a head-to-head trial

The manufacturer's submission references the EVIDENCE head to head trial between Avonex and Rebif-44 (EVIDENCE).[14] Unfortunately hazard ratios were not reported for this trial – hence the ERG has had to revert to using relative risks in the analysis here. An indirect comparison (Bucher et al., 1997) was applied to the results reported from the EVIDENCE trial in conjunction with results from the FREEDOMS and TRANSFORMS trials to derive relative risks of progression and relapse for Rebif-44 as compared to BSC. Results are shown in Table 39 below.[51]

Table 39: Relative Treatment Effects Rebif-44

	Population 1b	Population 1b but not 2
Relative Risk of Progression		
Rebif-44 vs. Placebo	[redacted]	[redacted]
Avonex vs. Placebo	[redacted]	[redacted]
Fingolimod vs. Placebo	[redacted]	[redacted]
Relative Risk of Relapse		
Rebif-44 vs. Placebo	[redacted]	[redacted]
Avonex vs. Placebo	[redacted]	[redacted]
fingolimod vs. Placebo	[redacted]	[redacted]

These relative effectiveness values were used in the model to calculate the deterministic cost-effectiveness results for the two populations; cost-effectiveness results are shown in the tables below. Note that the manufacturer's model does not allow for probabilistic results to be calculated for multiple comparators - hence deterministic results are reported here. While

these limitations make absolute values reported invalid, the ERG still considers the relative values to be informative. We can see from the results that while Rebif-44 dominates Avonex in both populations (is less expensive and more effective), it is still itself dominated by BSC or extendedly dominated by fingolimod in both populations.

Table 40:Incremental Analysis Including Rebif-44 (head to head trial) - Population 1b. Results are deterministic and relative risks are being used

	Total Cost (£)	Total QALYs	ICER (£ per QALY)
BSC	224,311	3.81	-
Rebif-44	258,458	4.13	ED (ICER of 107,701)
Avonex	271,646	3.98	D
fingolimod	321,730	4.88	**91,059**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 41: Incremental Analysis Including Rebif-44 (head to head trial) - Population 1b but not 2. Results are deterministic and relative risks are being used

	Total Cost (£)	Total QALYs	ICER (£ per QALY)
BSC	219,738	3.81	-
Rebif-44	261,437	3.44	D
Avonex	278,317	2.98	D
fingolimod	316,752	5.03	**79,315**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

(ii) adding Rebif as a comparator using evidence from the MTC

The ERG considered it useful to further explore the efficacy results using the relative effectiveness values from the manufacturer's MTC. The results of this analysis for the two populations considered are presented in the tables below. Using the data from the MTC analysis we see that in both populations Rebif-44 emerges as the main comparator, dominating Avonex as above, but not in this case being dominated or extendedly dominated itself.

Table 42: Incremental Analysis Including Rebif-44 (MTC) - Population 1b

	Total Cost (£)	Total QALYs	ICER (£ per QALY)
BSC	224,311	3.81	-
Rebif-44	255,735	4.29	66,322
Avonex	261,398	4.23	D
fingolimod	318,065	4.81	**119,213**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 43: Incremental Analysis Including Rebif-44 (MTC) - Population 1b but not 2

	Total Cost (£)	Total QALYs	ICER (£ per QALY)
BSC	219,738	3.81	-
Rebif-44	251,128	4.28	67,201
Avonex	256,733	4.22	D
fingolimod	313,122	4.79	**119,746**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

The results in this section suggest that Avonex is potentially dominated both by other beta-interferons as well as in some case by BSC. These findings cast further doubt upon the appropriateness of the sole use of Avonex as a comparator in the analysis.

6.4 Model Structure

A significant issue with the model structure is the way that treatment effects are applied to relapse rates. Relapse rates are modelled to be both dependent on progression and to be adjusted by the relative risk of relapse for being on a particular DMT as compared to BSC. These two adjustments, indirect due to progression and direct due to relative risk of relapse, are taken from different datasets and so have no implicit correlation; neither is this correlation explicitly dealt with in the model. The implication being that DMT impact on relapse is to some extent double counted in the model. To explore the full extent of the impact that this double counting could have on the model results, the ERG has re-run the model with all direct treatment effect adjustments to relapse rates excluded, leaving any impact on relapse rates being due only to indirect effects via the treatment impact on progression. The results of this analysis for the two populations considered are shown in the tables below.

Table 44: Only Indirect Treatment Effect on Relapse - Population 1b

	Total Cost (£)	Total QALYs	ICER (£ per QALY)
BSC	224,251	3.63	-
Avonex	273,072	3.72	ED (ICER of 537,603)
fingolimod	327,392	4.55	**112,294**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 45: Only Indirect Treatment Effect on Relapse - Population 1b but not 2

	Total Cost (£)	Total QALYs	ICER (£ per QALY)
BSC	219,399	4.68	-
Avonex	275,160	4.04	D
fingolimod	321,590	5.72	**98,019**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

It is apparent from the results in the tables above that the ICERs in both populations increase significantly in this scenario as compared to the base case results. While this scenario is an extreme case, it presents an upper bound on the impact of correcting for the structural limitations of the model in ignoring the correlation between relapse and progression.

6.5 Natural History

As was described in Section 5, analysis undertaken by the ERG showed that the underlying progression rates predicted by the model seem to over-estimate progression as observed in the trials. The manufacturer did not justify the divergence between model predictions and the trial observations in the submission. This section details the findings of the ERG on the impact of altering the natural history progression rates on model predictions.

Four scenarios have been implemented to examine the model's sensitivity to natural history progression rates: reduction in natural history progression transitions by 50%, 25% and 10%, as well as an increase in natural history progression transitions by 10%. The cost-effectiveness results are presented in Table 46 and Table 47, and show that reducing the natural history progression rates substantially increases the ICER for fingolimod – with a 50% reduction increasing the ICER to £252,147 in population 1b and to £191,027 in population 1b but not 2.

The model predictions appear to be very sensitive to the natural history progression data used. The plausibility of each of these scenarios needs to be carefully assessed.

Table 46: Modifying Natural History Progression Rates - Population 1b

	50% Reduction Natural History Progression Rate			25% Reduction Natural History Progression Rate			10% Reduction Natural History Progression Rate			10 % Increase Natural History Progression Rate		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	167,309	6.82	-	200,125	4.99	-	215,639	4.12	-	231,866	3.17	-
Avonex	228,533	6.59	D	254,179	4.96	D	266,270	4.15	ED (ICER of 1,773,946)	278,154	3.33	ED (ICER of 287,077)
fingolimod	285,624	7.29	**252,147**	306,819	5.81	**130,946**	316,724	5.07	**106,541**	326,732	4.26	**87,086**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 47: Modifying Natural History Progression Rates - Population 1b but not 2

	50% Reduction Natural History Progression Rate			25% Reduction Natural History Progression Rate			10% Reduction Natural History Progression Rate			10 % Increase Natural History Progression Rate		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	164,821	6.77	-	195,967	4.97	-	211,169	4.13	-	227,148	3.22	-
Avonex	234,350	5.91	D	259,065	4.21	D	270,686	3.45	D	281,353	2.73	D
fingolimod	282,386	7.38	**191,027**	302,178	5.94	**109,023**	311,652	5.26	**88,384**	321,394	4.53	**71,770**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

6.6 Treatment Effectiveness

In Section 5, we highlighted that despite other evaluations having used shorter time horizons, the use of a 50 year model time horizon in order to fully represent lifetime costs and effects associated with the alternative treatments is considered appropriate by the ERG. However, the ERG felt that it would be useful to evaluate the assumptions used in extrapolating treatment effects: the manufacturer's submission assumes that those patients receiving DMTs have a constant and continued treatment effect, so long as they remain on treatment, over the 50 year time horizon of the model. The ERG considers this extrapolation assumption for the treatment effect - informed by a 12 month trial and a 24 month trial - to be overly optimistic. In the manufacturer's sensitivity analysis, scenarios exploring possible waning of treatment effect were considered. This section details the ERG's extension of this work.

The scenarios presented here all assume full treatment efficacy initially, with this efficacy waning at either 2 years (the duration of the longest trial) or at 5 years. Treatment efficacy was modelled to wane to either 50%, 25% or 0% of the original level after the initial full efficacy period, with patients remaining on treatment throughout. A scenario where treatment efficacy dropped to 0% and treatment was discontinued immediately after the initial full efficacy period is also presented. The cost-effectiveness results of these scenarios for the two populations are shown in the four tables below.

In all scenarios, Avonex is either dominated or extendedly dominated and fingolimod is compared with BSC. In all cases ICERs are higher when waning occurs after 2 years than in the equivalent scenarios where waning occurs after 5 years. In all cases the more efficacy is reduced (waning to a lower percentage of efficacy) while remaining on treatment the higher the ICER. This is because the costs of treatment are still incurred but less health benefit is attained. Discontinuation of treatment (thus 0% efficacy) represents the lowest ICER scenario for any given treatment effect time horizon.

It is evident from these analyses that using more conservative extrapolation assumptions on the long term effectiveness of the alternative treatments can significantly increase the ICER for fingolimod (in some cases to more than double the base case results).

Table 48: Treatment Effect Waning at 2 years - Population 1b

	Waning at 2 years: reduction to 50% efficacy levels			Waning at 2 years: reduction to 25% efficacy levels			Waning at 2 years: reduction to 0% efficacy levels			Waning at 2 years: reduction to 0% efficacy levels; treatment discontinued		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	223,721	3.62	-	223,965	3.65	-	224,190	3.64	-	224,841	3.66	-
Avonex	273,465	3.58	D	274,238	3.55	D	274,922	3.50	D	243,852	3.63	D
fingolimod	322,781	4.33	**140,282**	323,579	4.21	**177,674**	324,405	4.04	**249,735**	261,507	3.98	**111,850**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 49: Treatment Effect Waning at 5 years - Population 1b

	Waning at 5 years: reduction to 50% efficacy levels			Waning at 5 years: reduction to 25% efficacy levels			Waning at 5 years: reduction to 0% efficacy levels			Waning at 5 years: reduction to 0% efficacy levels; treatment discontinued		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	224,422	3.64	-	222,994	3.63	-	224,026	3.62	-	224,435	3.64	-
Avonex	273,633	3.65	ED (ICER of 3,980,348)	272,373	3.61	D	274,050	3.55	D	256,863	3.64	D
fingolimod	322,709	4.50	**114,532**	321,680	4.38	**131,135**	322,837	4.31	**143,869**	287,374	4.28	**98,224**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 50: Treatment Effect Waning at 2 years - Population 1b but not 2

	Waning at 2 years: reduction to 50% efficacy levels			Waning at 2 years: reduction to 25% efficacy levels			Waning at 2 years: reduction to 0% efficacy levels			Waning at 2 years: reduction to 0% efficacy levels; treatment discontinued		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	219,424	3.63	-	220,371	3.64	-	219,640	3.63	-	219,493	3.63	-
Avonex	273,860	3.08	D	274,202	3.17	D	272,684	3.24	D	242,499	3.35	D
fingolimod	317,404	4.46	**117,439**	319,098	4.27	**156,282**	318,939	4.09	**216,863**	255,318	4.01	**93,662**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 51: Treatment Effect Waning at 5 years - Population 1b but not 2

	Waning at 5 years: reduction to 50% efficacy levels			Waning at 5 years: reduction to 25% efficacy levels			Waning at 5 years: reduction to 0% efficacy levels			Waning at 5 years: reduction to 0% efficacy levels; treatment discontinued		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	219,214	3.62	-	219,855	3.60	-	219,231	3.65	-	219,823	3.65	-
Avonex	274,200	3.01	D	274,460	3.05	D	273,765	3.11	D	258,880	3.18	D
fingolimod	316,496	4.60	**98,813**	317,346	4.50	**108,926**	317,018	4.44	**122,957**	281,199	4.40	**81,988**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

6.7 Health Related Quality of Life

The utility estimates used in the manufacturer's base case appear to be selected arbitrarily and the impact on model results of using alternative utility values has not been investigated. Utility data was collected from the manufacturer's trials, and was reported to the ERG in the clarifications from the manufacturer, but was not used in the model. Utility data used in previous NICE technology appraisals and from other external sources was also reported in the submission but not used. The ERG attempted to validate model predictions of baseline utility against those reported from the manufacturer's trials and found there to be significant differences in both the direction and magnitude between these values.

This section describes the ERG's attempt to explore scenarios using utility data from the manufacturer's trial in the model. There were a number of challenges involved in using this trial data: the manufacturer only provided values for patients in RRMS states, and only for patients in EDSS states 0 to 6. Following the assumptions used in the manufacturer's submission, a fixed decrement of 0.045 in utility was used to convert utility from RRMS to SPMS states. Missing values for EDSS states 7-9 were imputed using the relationship between EDSS states from either the data used in the model (Orme et al) or alternatively from the data used in a previous NICE technology appraisal for MS (TA 127); these alternative imputation assumptions constitute the two scenarios explored in this section.[5, 45] The utility values used and the data from which they were derived are presented in Table 52 below. Uncertainty estimates (e.g. variance or confidence intervals) were not provided and thus not considered in the cost-effectiveness analyses implemented in this section.

Table 52: Utilities Derivation for Scenario Analysis

	Orme	TA 127	FREEDOMS	TRANSFORMS	Trial Average	Trial Average + Orme	Trial Average + TA 127
EDSS							
0	0.870	0.959	[redacted]	[redacted]	0.895	0.895	0.895
1	0.799	0.688	[redacted]	[redacted]	0.870	0.870	0.870
2	0.705	0.688	[redacted]	[redacted]	0.840	0.840	0.840
3	0.574	0.645	[redacted]	[redacted]	0.755	0.755	0.755
4	0.610	0.61	[redacted]	[redacted]	0.690	0.690	0.690
5	0.518	0.581	[redacted]	[redacted]	0.640	0.640	0.640
6	0.460	0.538	[redacted]	[redacted]	0.545	0.545	0.545
7	0.297	0.410				0.382	0.417
8	-0.049	0.288				0.036	0.295
9	-0.195	0.049				-0.110	0.056

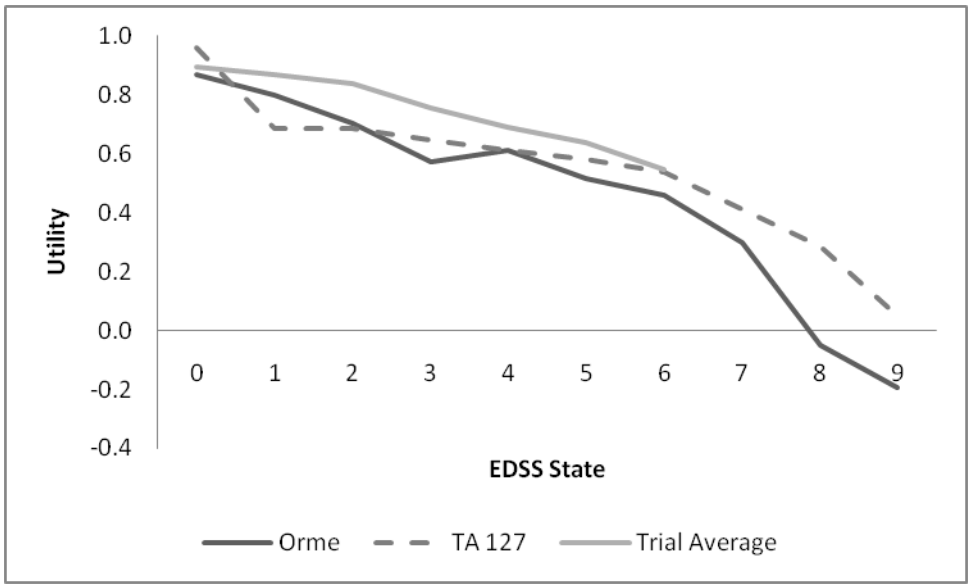

Figure 21: Patient Utility Data

Figure 21 above shows the three sources of utility data used in this analysis. The figure suggests that imputing the missing trial based utilities using TA 127 values would produce a less steep utility relationship than imputing with Orme values, thus representing a more conservative scenario. The cost-effectiveness results for these two scenarios using trial based utility data in the model are presented in the tables below.

Table 53: Utility Scenarios - Population 1b

	Orme et al. 2007 (Base Case)			Trial Average with imputed values from TA 127			Trial Average with imputed values from Orme et al 2007		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	224,192	3.66	-	224,085	7.68	-	223,589	5.71	-
Avonex	272,454	3.76	ED (ICER of 471,431)	272,264	7.75	ED (ICER of 720,874)	272,181	5.81	ED (ICER of 488,209)
fingolimod	321,995	4.70	**94,094**	321,863	8.60	**106,824**	321,421	6.81	**89,030**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Table 54: Utility Scenarios - Population 1b but not 2

	Orme et al. 2007 (Base Case)			Trial Average with imputed values from TA 127			Trial Average with imputed values from Orme et al 2007		
Treatment	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**	**Total Cost (£)**	**Total QALYs**	**ICER (£ per QALY)**
BSC	219,865	3.64	-	219,557	7.61	-	219,630	5.68	-
Avonex	274,611	3.06	D	275,410	6.97	D	275,206	4.98	D
fingolimod	316,649	4.83	**81,369**	316,280	8.66	**91,345**	316,312	6.95	**76,367**

D = Option ruled out by dominance (more expensive and less effective); ED = option ruled out by extended dominance

Looking at the results we see that changing the utility values of only three of the EDSS states has a significant impact on the ICER for fingolimod. While the model is highly sensitive to small changes in these values, there is no clear justification for the utility data selected by the manufacturer.

6.8 Summary

The ERG's additional exploratory analysis has shown the sensitivity of the manufacturer's model to alternative sources of parameter data and alternative modelling assumptions. While the data sources selected and assumptions made have not been adequately justified by the manufacturer, the ERG has established that alternative choices of these lead to significant differences in the cost-effectiveness results estimated. In particular the ERG has shown that estimates of cost-effectiveness results are highly sensitive to changes in: the initial EDSS population distribution, interventions and comparators, natural history progression rates, waning of treatment effect, utility estimates, and the way effectiveness on relapse rates has been dealt with within the submission. This was observed for both the populations analysed: population 1b and population 1b but not 2.

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

The manufacturer's submission rested primarily on the evidence from two well-conducted and adequately powered trials, TRANSFORMS and FREEDOMS which enrolled patients with RRMS. TRANSFORMS compared fingolimod 0.5 mg with fingolimod 1.25 mg (which is disregarded for the purposes of this appraisal) and with interferon- beta-1a (Avonex) 30 mg. FREEDOMS compared fingolimod at 0.5 mg and 1.25 mg with placebo. The primary outcome in both trials was annualised relapse rate (ARR). Disability progression, adverse events, MRI outcomes and health-related QoL were also reported. The duration of TRANSFORMS was 12 months; the duration of FREEDOMS was 24 months.

In both cases the trial populations were broader than that approved in the positive CHMP opinion. The manufacturer therefore defined post-hoc subgroups in both trials which approximated the CHMP indicated populations and designated one of these, population 1b, as the base case for the submission. Population 1b consisted of patients who were previously treated and have had equal or more relapses in year one than in year two.

Population 1b constituted fewer than 50% of the population of TRANSFORMS and fewer than 20% of the population of FREEDOMS; this subgroup showed a high degree of overlap with others, and a substantial number of patients who met criteria for population 2 in the CHMP indication (patients with RES RRMS) were included in population 1b. Since these patients were eligible for treatment with natalizumab, the ERG was concerned that including them in comparisons between fingolimod and Avonex or placebo was potentially inappropriate. The ERG therefore requested and obtained data on population 1b but not 2; this was used to explore alternatives to the manufacturer's base case in the economic model.

The results of the FREEDOMS and TRANSFORMS trials indicate that there is a benefit to fingolimod 0.5 mg compared to placebo and to Avonex on the outcome of the ARR for both population 1b and population 1b not 2, although there were no differences in disability progression. Serious adverse events were rare and were broadly comparable between the arms in both trials, and, while there were some differences in the incidence of specific adverse events these generally followed a predicted pattern (for instance a lower incidence of influenza-type illness in fingolimod-treated patients compared to those in the Avonex arm of TRANSFORMS).

The ERG's concern as to the appropriateness of Avonex as the comparator for both the clinical and cost-effectiveness data was not limited to patients meeting criteria for population 2. The approximation of population 1b, which the manufacturer presented as the base case, were patients who had failed to respond to prior treatment with DMT which, in the great majority of cases, was treatment with an interferon. In this population, therefore, fingolimod was compared to a treatment which was highly likely to be ineffective. Additional concern about the use of Avonex as the comparator in the submission arise from its effectiveness relative to the alternative formulations of beta-interferon-1a (Rebif) and beta-interferon-1b (Betaferon). Evidence from RCTs which have assessed head-to-head comparisons of the different interferon products, and the views of the ERG's clinical expert, suggested that Avonex may be the least effective of the available interferons.

Whilst the manufacturer's submission included an MTC in the clinical effectiveness section, this was not subsequently used to inform the economic model. Therefore the *de novo* model submitted by the manufacturer did not consider any comparator other than Avonex. The decision not to use the MTC was reasonable based on the level of heterogeneity between the trials and the fact that the data were based on general RRMS populations. However, in place of the results of the MTC the manufacturer presented an indirect comparison which showed that, for the outcome of disability progression, Avonex [redacted] placebo. The justification given for the use of Avonex as the main treatment comparator to fingolimod was the availability of relevant data.

There is therefore considerable uncertainty surrounding the efficacy of fingolimod in the base case population relative to an appropriate compatator. The relationship between the placebo arm of the FREEDOMS trial and BSC is unclear; the ERG's clinical advisor has indicated that, for patients with low EDSS scores, the placebo arm of a clinical trial may provide a more intensive level of medical attention than would be the case with BSC in the community. However, the number of patients in the FREEDOMS trial who met the criteria for the post-hoc approximation of population 1b was low, at 19.7% of the randomised population. A third of the trial population were randomised to fingolimod 1.25 mg, and a further proportion of patients additionally met criteria for population 2. The numbers of patients in this post-hoc subgroup on which the estimate of relative efficacy compared to a placebo approximation of BSC is based are therefore small, and must be subject to a high degree of uncertainty.

7.2 Summary of cost-effectiveness issues

The manufacturer does not appear to have used a systematic approach to identify and select appropriate data sources to inform the key parameters of the model – choices of data appear to be arbitrary and unjustified. Methods used for subsequently deriving the various model parameters from the selected data are not fully described and assumptions made in using these methods are not discussed or justified. There has been no attempt by the manufacturer to validate the predictions of the model either internally against the trial data or externally against other published studies or clinician opinion.

To explore the impact of the values of key parameters and the assumptions used by the manufacturer, the ERG evaluated a number of scenarios using alternative sources of evidence referred to in the manufacturer's submission but not used in the model and using alternative modelling assumptions. This additional analysis demonstrated that cost-effectiveness results produced by the manufacturer's model are highly sensitive to changes in: the initial EDSS population distribution, interventions and comparators, natural history progression rates, waning of treatment effect, utility estimates, and the way effectiveness on relapse rates has been dealt with within the submission. This was observed for both population 1b and population 1b but not 2. In all the scenarios explored, Avonex was either dominated or extendedly dominated by BSC. None of the scenarios explored suggested fingolimod to be cost-effective at the usual NICE thresholds.

7.3 Implications for research

Given the uncertainties surrounding the efficacy of fingolimod in the CHMP indicated populations, and the concerns over the use of Avonex as the main comparator, there is a need for head-to-head trials of fingolimod 0.5 versus natalizumab in the RES population (population 2 of the CHMP indication). There is also a need for a trial comparing fingolimod with true BSC in population 1 of the indication; this should include patients who meet the criteria for population 1a, in addition to those who meet criteria for 1b.

8 REFERENCES

1. Kappos L, Radue E-W, O'Connor P, Polman CH, Hohlfeld R, Calabresi P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* 2010;362:387-401.
2. Cohen JA, Barkhof F, Comi G, Hartung H-P, Khatri BO, Montalban X, et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N Engl J Med* 2010;362:402-15.
3. Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L, Miller DH, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med* 2006;354:899-910.
4. National Institute for Health and Clinical Excellence. *Beta interferon and glatiramer acetate for the treatment of multiple sclerosis. Technology Appraisal No. 32*. NICE; 2002. [cited 2010 Apr 22]. Available from: http://www.nice.org.uk/nicemedia/live/11441/32290/32290.pdf.
5. National Institute for Health and Clinical Excellence. *Natalizumab for the treatment of adults with highly active relapsing–remitting multiple sclerosis [TA127]*. NICE; 2007. [cited 2010 Apr 22]. Available from: http://www.nice.org.uk/nicemedia/live/11441/32290/32290.pdf.
6. Pokorski RJ. Long-term survival experience of patients with multiple sclerosis. *J Insur Med* 1997;29:100-05.
7. Sadovnick AD, Ebers GC, Wilson RW, Paty DW. Life expectancy in patients attending multiple sclerosis clinics. *Neurology* 1992;42:991-4.
8. 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.
9. National Institute For Health And Clinical Excellence. *Natalizumab for the treatment of adults with highly active relapsing–remitting multiple sclerosis*. London: Nice; 2007.
10. Vartanian T. An examination of the results of the EVIDENCE, INCOMIN and phase III studies of interferon beta products in the treatment of multiple sclerosis. *Clin Ther* 2002;25:105-18.
11. Nikfar S, Rahimi R, Abdollahi M. A systematic review on the efficacy of interferon beta in relapsing remitting multiple sclerosis: comparison of different formulations. *Int J Pharmacol* 2010;6:638-44.
12. European Medicines Agency. *Assessment report: Gilenya*. London: EMA; 2011.
13. Rice GPA, Incorvaia B, Munari LM, Ebers G, Polman C, D'Amico R, et al. Interferon in relapsing-remitting multiple sclerosis. *Cochrane Database Syst Rev 2001, Issue 4. Art. No.: CD002002. DOI: 10.1002/14651858.CD002002.*
14. Panitch H, Goodin DS, Francis G, Chang P, Coyle PK, O'Connor P, et al. Randomized, comparative study of interferon beta-1a treatment regimens in MS. The EVIDENCE Trial. *Neurology* 2002;59:1496-506.
15. Etemadifar M, Janghorbani M, Shaygannejad V. Comparison of Betaferon, Avonex, and Rebif in treatment of relapsing-remitting multiple sclerosis. *Acta Neurol Scand* 2006;113:283-7.
16. Limmroth V, Malessa R, Zettl UK, Koehler J, Japp G, Haller P, et al. Quality Assessment in Multiple Sclerosis Therapy (QUASIMS): a comparison of interferon beta therapies for relapsing-remitting multiple sclerosis. *J Neurol* 2007;254:67-77.
17. Khan OA, Tselis AC, Kamholz JA, Garbern JY, Lewis RA, Lisak RP. A prospective, open-label treatment trial to compare the effect of IFNb-1a (Avonex), IFNb-1b (Betaseron), and glatiramer acetate (Copaxone) on the relapse rate in relapsing-remitting multiple sclerosis: results after 18 months of therapy. *Mult Scler* 2001;7:349-53.
18. Patti F, Pappalardo A, Florio C, Politi G, Fiorilla T, Reggio E, et al. Effects of interferon beta-1a and -1b over time: 6-year results of an observational head-to-head study. *Acta Neurol Scand* 2006;113:241-47.
19. Kappos L, Antel J, Comi G, Montalban X, O'Connor P, Polman CH, et al. Oral fingolimod (FTY720) for relapsing multiple sclerosis. *N Engl J Med* 2006;355:1124-40.
20. Comi G, O'Connor P, Montalban X, Antel J, Radue E-W, Karlsson G, et al. Phase II study of oral fingolimod (FTY720) in multiple sclerosis: 3-year results. *Mult Scler* 2010;16:197-207.

21. O'Connor P, Comi G, Montalban X, Antel J, Radue EW, de Vera A, et al. Oral fingolimod (FTY720) in multiple sclerosis: two-year results of a phase II extension study *Neurology* 2009;72:73-79.
22. Durelli L, Verdun E, Barbero P, Bergui M, Versino E, Ghezzi A, et al. Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: results of a 2-year prospective randomised multicentre study (INCOMIN). *Lancet* 2002;359:1453-60.
23. Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, et al. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Ann Neurol* 1996;39:285-94.
24. The IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 1993;43:655-61.
25. PRISMS Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. *Lancet* 1998;352:1498-504.
26. O'Connor P, Filippi M, Arnason B, Comi G, Cook S, Goodin D, et al. 250 microg or 500 microg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study. *Lancet Neurology* 2009;8:889-97.
27. Cadavid D, Wolansky LJ, Skurnick J, Lincoln J, Cheriyan J, Szczepanowski K, et al. Efficacy of treatment of MS with IFNbeta-1b or glatiramer acetate by monthly brain MRI in the BECOME study. *Neurology* 2009;72:1976-83.
28. Mikol DD, Barkhof F, Chang P, Coyle PK, Jeffery DR, Schwid SR, et al. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REbif vs Glatiramer Acetate in Relapsing MS Disease[REGARD] study): a multicentre, randomised, parallel, open-label trial. *Lancet Neurology* 2008;7:903-14.
29. Hurwitz BJ, Jeffery D, Arnason B, Bigley K, Coyle P, Goodin D, et al. Tolerability and safety profile of 12- to 28-week treatment with interferon beta-1b 250 and 500 microg QOD in patients with relapsing-remitting multiple sclerosis: a multicenter, randomized, double-blind, parallel-group pilot study. *Clin Ther* 2008;30:1102-12.
30. Wroe SJ. Effects of dose titration on tolerability and efficacy of interferon beta-1b in people with multiple sclerosis. *J Int Med Res* 2005;33:309-18.
31. Saida T, Tashiro K, Itoyama Y, Sato T, Ohashi Y, Zhao Z. Interferon beta-1b is effective in Japanese RRMS patients: a randomized, multicenter study. *Neurology* 2005;64:621-30.
32. Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. *Neurology* 1995;45:1268-76.
33. Comi G, Filippi M, Wolinsky JS. European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging–measured disease activity and burden in patients with relapsing multiple sclerosis. *Ann Neurol* 2001;49:290–97.
34. Bornstein MB, Miller A, Slagle S, Weitzman M, Crystal H, Drexler E, et al. A pilot trial of Cop 1 in exacerbating-remitting multiple sclerosis. *N Engl J Med* 1987;317:408-14.
35. Ford HL, Gerry E, Tennant A, Whalley D, Haigh R, Johnson MH. Developing a disease-specific quality of life measure for people with multiple sclerosis. *Clin Rehabil* 2001;15:247-58.
36. Eckermann S, Coory M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. *J Clin Epidemiol* 2009;62:1031-6.
37. 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.
38. Prosser LA, Kuntz KM, Bar-Or A, Weinstein MC. Patient and community preferences for treatments and health states in multiple sclerosis. *Mult Scler* 2003;9:311-9.
39. Department of Health. *NHS reference costs 2009-2010: national schedule of reference costs, year: "2009-10". NHS Trusts and PCTs combined*. DoH; 2011. [cited 2011 Feb 25]. Available from:

http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123459.
40. Joint Formulary Committee. *British National Formulary*. 60th ed. London: Pharmaceutical Press; 2010.
41. Association of British Neurologists. *Revised (2009) guidelines for prescribing in multiple sclerosis*. London: Association of British Neurologists; 2009 Nov.
42. Peninsula Technology Assessment Group (PenTAG). *The effectiveness and cost-effectiveness of natalizumab for multiple sclerosis: an evidence review of the submission from Biogen*. Exeter: Institute of Health and Social Care Research at the Peninsula Medical School; 2007.
43. Office for National Statistics (ONS). *Interim life tables, England and Wales, 1980-82 to 2007-09*. ONS; 2010. [cited 2010 Oct 12]. Available from: http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=14459.
44. Weinshenker BG, Bass B, Rice GP, Noseworthy J, Carriere W, Baskerville J, et al. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. *Brain* 1989;112:133-46.
45. 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.
46. 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.
47. Gani R, Nixon RM, Hughes S, Jackson CH. Estimating the rates of disability progression in people with active relapsing-remitting multiple sclerosis. *J Med Econ* 2007;10:79-89
48. 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.
49. Biogen Idec Limited. *Avonex: summary of product characteristics*. Biogen Idec Limited; 2009. [cited 2010 Apr 23]. Available from: http://www.medicines.org.uk/EMC/medicine/257/SPC/AVONEX+30+micrograms+0.5+ml+solution+for+injection.
50. Curtis L. *Unit costs of health and social care 2010*. Personal Social Services Research Unit, The University of Kent; 2010. [cited 2011 Feb 25]. Available from: http://www.education.gov.uk/publications//eOrderingDownload/PSSRU-1368-230X.pdf.
51. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50:683-91.
52. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. *J Health Econ* 1993;12:459-67.

Appendix 1: Quality Assessment of the Economic Model

Quality criterion	***Question(s)***	***Response* (Y, N, or NS)**	***Comments***
S1	Is there a clear statement of the decision problem?	*?*	Sub-population issue.
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?	*?*	
	Is the primary decision-maker specified?	*Y*	The report is written for NICE
S2	Is the perspective of the model stated clearly?	*Y*	NHS and PSS perspective is taken, caregiver disutilities are also considered
	Are the model inputs consistent with the stated perspective?	*Y*	
	Has the scope of the model been stated and justified?	*?*	
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	*Y*	The model measures progression of MS (based on disability progression, relapse rate, mortality rate and adverse events)
S3	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	*N*	While it appears consistent with previous models, it does not seem coherent with the clinical observations of progression resulting as a consequence of relapse.
	Are the sources of data used to develop the structure of the model specified?	*Y*	
	Are the causal relationships described by the model structure justified appropriately?	*N*	Issue of potential correlation between relapse and progression is overlooked. In addition to the connection of progression and mortality.
S4	Are the structural assumptions transparent and justified?	*N*	50 year time horizon used compared to 20 years in previous models, no justification given
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	*?*	See above *Comparison with previous models*

S5	Is there a clear definition of the options under evaluation?	*Y*	The treatments not considered were not well justified
	Have all feasible and practical options been evaluated?	*N*	Other treatments have been overlooked as only results for Avonex are presented
	Is there justification for the exclusion of feasible options?	*Y*	However justification presented is not convincing
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	*Y*	Though key correlations are not captured by the model
S7	Is the time horizon of the model sufficient to reflect all important differences between options?	*Y*	
	Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	*N*	Duration of treatment and treatment effect is not well justified comparison needed with the trial data and other models
S8	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	*N*	Questions have been raised in clinician consultations whether RRMS and SPMS progressed in the manner modelled.
S9	Is the cycle length defined and justified in terms of the natural history of disease?	*Y*	Cycle length is presented but not justified
D1	Are the data identification methods transparent and appropriate given the objectives of the model?	*N*	Lack of clarity about where data originates from and how it is converted for use in the model
	Where choices have been made between data sources, are these justified appropriately?	*N*	Little justification is given when not using FREEDOMS or TRANSFORMS studies, or implications
	Has particular attention been paid to identifying data for the important parameters in the model?	*N*	No systematic review was conducted to identify important parameters
	Has the quality of the data been assessed appropriately?	*N*	
	Where expert opinion has been used, are the methods described and justified?	*N*	Experts were consulted but no formal results are presented

D2	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	*N*	Modelling assumptions used not justified
D2a	Is the choice of baseline data described and justified?	*N*	Sub-population and comparator issues, lack of justification for use of Ontario, no justification of assumptions linking BSC with placebo (as used in head to head trials)
	Are transition probabilities calculated appropriately?	*N*	Lack of clarity about calculations
	Has a half-cycle correction been applied to both cost and outcome?	*Y*	
	If not, has this omission been justified?	*NA*	
D2b	If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	*Y*	Tough much potentially relevant data is ignored, for example, the MTC conducted as part of the clinical effectiveness review was not used.
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	*N*	No justification for the extrapolation of short term trial data over 50 years is given. This is an important omission since the modelled time horizon is significantly longer than other previous NICE appraisals for DMTs.
	Have alternative extrapolation assumptions been explored through sensitivity analysis?	*Y*	
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	*NA*	
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	*Y*	
D2c	Are the costs incorporated into the model justified?	*Y*	
	Has the source for all costs been described?	*Y*	
	Have discount rates been described and justified given the target decision-maker?	*Y*	Use of 3.5% for cost and utilities as is consistent with NICE reference
D2d	Are the utilities incorporated into the model appropriate?	*Y*	Dependent on the assumption of consistency of EDSS states holding and the Orme et al. (2007) study being appropriate

	Is the source for the utility weights referenced?	*Y*	Use of Orme et al. (2007)
	Are the methods of derivation for the utility weights justified?	*Y*	Simple application of reference
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	*N*	Lack of certainty behind much of the data, including Ontario data and Patzold and Pocklington (1982)
	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	*N*	No attempt was made in the submission to validate the model with external data
	Is the process of data incorporation transparent?	*N*	Data used for model is given in submission but not always where it came from or how it is calculated
	If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	*N*	Table 58 gives list of variables, there appears to be data that could have a distribution that has been deemed deterministic
	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	*N*	It is not clear what second order uncertainties are considered
D4	Have the four principal types of uncertainty been addressed?	*N*	Structural, methodological, heterogeneity and parameter
	If not, has the omission of particular forms of uncertainty been justified?	*N*	
D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	*N*	
D4b	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	*Y*	Not addressed in a comprehensive manner
D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	*N*	No sub-group analysis was conducted
D4d	Are the methods of assessment of parameter uncertainty appropriate?	*Y*	PSA methods are used

	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	*N*	The distribution of variables and ranges used in table 58, 79 and 80 respectively seem to be largely arbitrary. Unsure where CIs given in table 79 come from, e.g. for utilities Orme et al (2007) doesn't give CI so do they come from other references?
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	*Y*	Verification of the model in recorded in table 81
C2	Are any counterintuitive results from the model explained and justified?	*N*	No counter-intuitive results are reported in verification, however, the issue of why comparison of fingolimod to BSC results in a very large ICER is unclear.
	If the model has been calibrated against independent data, have any differences been explained and justified?	*NA*	
	Have the results of the model been compared with those of previous models and any differences in results explained?	N	Models are not compared

Appendix 2 – Conversion Calculations

Conversions from RRMS to SPMS were also calculated using the London Ontario dataset. The median time to conversion observed in the data was used to define an exponential distribution for EDSS 1 parameterized using the relation:

$$m[X] = \frac{ln[2]}{\lambda}$$

where λ is the parameter of the exponential distribution and m[X] is median time to conversion to SPMS when in RRMS EDSS state X.

A Cox proportional hazards model was fitted to EDSS as a continuous value, and was used to define the relationship between covariates (EDSS is this case) and the hazard ratio of conversion to SPMS between the base case EDSS 1 and all other EDSS states.

$$ln[\frac{H(t)}{H(t)_{DSS1}}] = \beta \times X$$

This provided the relationship (β) between EDSS score and the log of the hazard ratio (between the hazard rate of EDSS X and EDSS 1). The annual probability of conversion to SPMS given that the subject was in EDSS state 1 was calculated using the following equation:

$$P = 1 - e^{(-\lambda)}$$

To estimate the probability of conversion for all other EDSS states the hazard rates were calculated based on:

$$H(t) = \lambda \times e^{\beta \times X}$$

Where X is the EDSS state, and ☐ is the hazard rate for conversion from EDSS 1.

Appendix 3 – Progression Calculations

Table 55 and Table 56 below show the progression data from the FREEDOMS and TRANSFORMS trials respectively. In these tables **n** represents the number of patients progressing, **N** represents the overall patient population and the risk of progression is given as the ratio of the two. The calculations that follow show how the trial data are used to derive the relative risks used in the model.

Table 55: Progression Data FREEDOMS Trial (extracted from manufacturer's model)

	n	N	Risk (n/N)
Placebo	[redacted]	[redacted]	[redacted]
fingolimod	[redacted]	[redacted]	[redacted]

RR_{FP} = Risk(fingolimod)/Risk(Placebo) = [redacted]

RR_{PF} = Risk(Placebo)/Risk(fingolimod) = [redacted]

Var (ln RR_{FP}) = Var (ln RR_{PF}) = [1-Risk(Placebo)]/n(Placebo) + [1-Risk(fingolimod)]/n(fingolimod) = [redacted]

$LCI_{FP} = e^{[\ln RR_{FP} - 1.96 * \sqrt{Var(\ln RR_{FP})}]}$ = [redacted]

$UCI_{FP} = e^{[\ln RR_{FP} + 1.96 * \sqrt{Var(\ln RR_{FP})}]}$ = [redacted]

Table 56: Progression Data TRANSFORMS Trial (extracted from manufacturer's model)

	n	N	Risk (n/N)
fingolimod	[redacted]	[redacted]	[redacted]
Avonex	[redacted]	[redacted]	[redacted]

RR_{AF} = Risk(Avonex)/Risk(fingolimod) = [redacted]

Var (ln RR_{AF}) = [1-Risk(Avonex)]/n(Avonex) + [1-Risk(fingolimod)]/n(fingolimod) = [redacted]

$RR_{AP} = e^{[\ln RR_{AF} - \ln RR_{PF}]}$ = [redacted]

Var (ln RR_{AP}) = Var (ln RR_{AF}) + Var (ln RR_{PF}) = [redacted]

$LCI_{FP} = e^{[\ln RR_{AP} - 1.96 * \sqrt{Var(\ln RR_{AP})}]}$ = [redacted]

$UCI_{FP} = e^{[\ln RR_{AP} + 1.96 * \sqrt{Var(\ln RR_{AP})}]}$ = [redacted]

Appendix 4 – Relapse Calculations

Table 57 and Table 58 below show the relapse data from the FREEDOMS and TRANSFORMS trials respectively. In these tables **n** represents the number of relapses observed, **N** represents the overall patient population and the risk of relapse is given as the ratio of the two. The calculations that follow show how the trial data are used to derive the relative risks used in the model.

Table 57: Relapse Data FREEDOMS Trial (extracted from manufacturer's model)

	n	N	Risk (n/N)
Placebo	[redacted]	[redacted]	[redacted]
fingolimod	[redacted]	[redacted]	[redacted]

RR_{FP} = Risk(fingolimod)/Risk(Placebo) = [redacted]

RR_{PF} = Risk(Placebo)/Risk(fingolimod) = [redacted]

Var (ln RR_{FP}) = Var (ln RR_{PF}) = [1-Risk(Placebo)]/n(Placebo) + [1-Risk(fingolimod)]/n(fingolimod) = [redacted]

$LCI_{FP} = e^{[\ln RR_{FP} - 1.96 * \sqrt{Var(\ln RR_{FP})}]}$ = [redacted]

$UCI_{FP} = e^{[\ln RR_{FP} + 1.96 * \sqrt{Var(\ln RR_{FP})}]}$ = [redacted]

Table 58: Relapse Data TRANSFORMS Trial (extracted from manufacturer's model)

	n	N	Risk (n/N)
fingolimod	[redacted]	[redacted]	[redacted]
Avonex	[redacted]	[redacted]	[redacted]

RR_{AF} = Risk(Avonex)/Risk(fingolimod) = [redacted]

Var (ln RR_{AF}) = [1-Risk(Avonex)]/n(Avonex) + [1-Risk(fingolimod)]/n(fingolimod) = [redacted]

$RR_{AP} = e^{[\ln RR_{AF} - \ln RR_{PF}]}$ = [redacted]

Var (ln RR_{AP}) = Var (ln RR_{AF}) + Var (ln RR_{PF}) = [redacted]

$LCI_{AP} = e^{[\ln RR_{AP} - 1.96 * \sqrt{Var(\ln RR_{AP})}]}$ = [redacted]

$UCI_{AP} = e^{[\ln RR_{AP} + 1.96 * \sqrt{Var(\ln RR_{AP})}]}$ = [redacted]