Evidence Review Group's Report
Dimethyl fumarate for treating relapsing-remitting multiple sclerosis

Produced by CRD and CHE Technology Assessment Group

Authors Gill Norman, Research Fellow, CRD
Stephen Rice, Research Fellow, CRD
Joanne O'Connor, Research Fellow, CRD
Kate Lewis-Light, Information Specialist, CRD
Dawn Craig, Research Fellow, CRD
Catriona McDaid, Senior Research Fellow, CRD

Correspondence to Dr Catriona McDaid
Centre for Reviews and Dissemination
University of York
York
YO10 5DD

Date completed 24/07/2013

Source of funding

This report was commissioned by the NIHR HTA Programme as project number 12/81/01.

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.

We would also like to thank Professor Stephen Palmer (Professor of Health Economics, CHE) for advice and for providing comments on the ERG report.

Rider on responsibility for report

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

This report should be referenced as follows:

Norman G, Rice S, O'Connor J, Lewis-Light K, Craig D, McDaid C. Dimethyl fumarate for treating relapsing-remitting multiple sclerosis: A Single Technology Appraisal, *CRD and CHE Technology Assessment Group*, 2013.

Contributions of authors:

Gill Norman wrote the clinical effectiveness sections of the report; Stephen Rice and Joanne O'Connor wrote the economic sections of the report; Kate Lewis-Light wrote the sections of the report dealing with search strategies, undertook searches for the ERG and provided information support. Dawn Craig and Catriona McDaid managed the cost-effectiveness and clinical effectiveness sections of the project respectively and reviewed and commented on the report.

.

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

ABN	Association of British Neurologists
ALT	Alanine transaminase
ARR	Annualised relapse rate
BID	Twice daily
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CSR	Clinical Study Report
DF	Dimethyl fumarate
DMT	Disease modifying therapy
EDSS	Expanded Disability Status Scale
EMA	European Medicines Agency
EOD	Every other day
EQ-5D	European Quality of Life 5 dimensions
FDA	(US) Food and Drug Administration
GI	Gastrointestional
HR	Hazard ratio
HRQoL	Health related quality of life
ICER	Incremental cost-effectiveness ratio
IM	Intramuscular
INEC	Independent neurology examination committee
ITT	Intention-to-treat
IV	Intravenous
KOL	Key opinion leader
MSM	Multi-state Markov
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MSFC	Multiple Sclerosis Functional Composite
MTC	Mixed treatment comparison

NA	Not available
NHS	National health service
OD	Once daily
PAS	Patient Access Scheme
PML	Progressive multifocal leukoencephalopathy
PO	Oral
PPMS	Primary progressive multiple sclerosis
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RES	Rapidly evolving severe
RR	Relative risk
RRMS	Relapsing remitting multiple sclerosis
RSS	Risk sharing scheme
SC	Subcutaneous
ScHARR	School of Health and Related Research
SD	Standard deviation
SF-36	Short-form 36
SPC	Summary of product characteristics
SPMS	Secondary-progressive multiple sclerosis
SUR	Seemingly unrelated regression
TIW	Three times weekly
TID	Three times daily
VAS	Visual analogue scale
WMD	Weighted mean difference

1 Summary

1.1 Critique of the decision problem in the manufacturer's submission

This report presents the evidence review group (ERG)'s assessment of the manufacturer's (Biogen) submission to NICE on the use of dimethyl fumarate (Tecfidera™), an oral drug for the treatment of relapsing remitting multiple sclerosis (RRMS) in adults (aged ≥ 18 years).

The remit of the NICE scope was to appraise dimethyl fumarate within its licensed indication for the treatment of relapsing remitting multiple sclerosis (RRMS) in adults (aged ≥ 18 years). Dimethyl fumarate does not currently have a UK marketing authorisation for the treatment of RRMS; although it has received a Committee for Medicinal Products for Human Use (CHMP) positive opinion for use in "adults patients with relapsing remitting multiple sclerosis". The scope outlines relevant comparators as beta-interferon, glatiramer acetate, natalizumab (for patients with rapidly evolving severe (RES) RRMS and fingolimod (for patients with highly active RRMS who have received treatment with beta-interferon). The overall population, the intervention and the outcomes in the manufacturer's submission are consistent with the National Institute for Health and Care Excellence (NICE) scope. However there is some inconsistency in the submission regarding the populations for which the comparators are evaluated. Natalizumab and fingolimod were evaluated in broad RRMS populations but the licensing and guidance recommendations were based on subgroup data. The trials of fingolimod and natalizumab were included in the manufacturer's mixed treatment comparison (MTC) for all RRMS patients which, given the trial populations, was relevant. However, licensing and NICE recommendations for these drugs were based on subgroup analysis. It is therefore not clear that the decision problems for these comparators (i.e. their effectiveness and cost-effectiveness compared to dimethyl fumarate in the appropriate subgroups) have been fully addressed. The manufacturer stated that this was due to lack of data on the effectiveness of interventions in the relevant subgroups. An additional comparator, the novel oral agent teriflunomide, was also included as a comparator in the MTC but not in the decision model.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

The manufacturer's submission centred on the evidence from two phase three randomised controlled trials (RCTs) which compared dimethyl fumarate (Tecfidera™) in its approved dose of 240mg orally twice daily (BID) with placebo: the DEFINE (N = 1,237) and CONFIRM (N = 1,430) trials. The CONFIRM study also used an active comparator, glatiramer acetate at the approved dose of 20 mg given once daily by subcutaneous (sc) injection which was included in the submission. The duration of both studies was 96 weeks. Both trials also included an arm treated with a higher dose of dimethyl fumarate of 240 mg three times daily (TID). The primary outcome in DEFINE was the proportion of patients experiencing relapse by two years; the primary outcome in CONFIRM was annualised relapse rate (ARR). Both trials reported disability progression sustained for three months on the

Expanded Disability Status Scale (EDSS) at 24 months as a secondary outcome. Disability progression sustained for six months was reported as a prespecified sensitivity analysis in both trials; six months sustained progression is regarded as a more robust measure of permanent disability progression.

A pooled analysis of the direct comparisons of dimethyl fumarate versus placebo from the DEFINE and CONFIRM studies was also presented for outcomes including ARR and EDSS disability progression. An MTC comprising a total of 27 RCTs was used to compare dimethyl fumarate with the active comparators defined in the NICE scope (forms of interferon beta-1a, interferon beta-1b, glatiramer acetate, natalizumab, and fingolimod); teriflunomide was also included.

ARR showed a statistically significant benefit of dimethyl fumarate compared to placebo in both the DEFINE (rate ratio 0.47, 95% CI 0.37 to 0.61) and CONFIRM trials (rate ratio 0.56, 95% CI 0.42 to 0.74). The primary outcome of DEFINE (proportion of patients experiencing relapse by two years) also showed a statistically significant benefit. Benefits of dimethyl fumarate compared to placebo were reflected in the pooled estimate of effect from the meta-analysis [redacted] The MTC showed a statistically significant benefit for dimethyl fumarate versus all comparators (all interferon beta medications, glatiramer acetate, teriflunomide and placebo) except natalizumab and fingolimod. There was a statistically significant benefit in favour of natalizumab and a non-significant effect in favour of fingolimod (Table 1).

Disability progression confirmed for at least three months showed a statistically significant benefit of dimethyl fumarate compared to placebo in the DEFINE study (HR 0.62, 95% CI 0.44 to 0.87); in the CONFIRM study the confidence intervals included the possibility of no benefit (HR 0.79, 95% CI 0.52 to 1.19). A pooled analysis of placebo comparisons from the DEFINE and CONFIRM trials showed statistically significant benefits for this outcome [redacted] Disability progression confirmed for at least six months showed less clear evidence of a benefit for dimethyl fumarate; a statistically significant benefit was seen only in the pooled analysis of DEFINE and CONFIRM (HR 0.71, 95% CI 0.52 to 0.97) (Table 2).

The MTC found that there were no statistically significant differences between dimethyl fumarate and any active comparator for three month confirmed disability progression; there was a statistically significant benefit for dimethyl fumarate compared to placebo for this outcome. Directions of effect favoured dimethyl fumarate for comparisons with the beta-interferons and glatiramer acetate but not natalizumab and fingolimod. The analysis for progression confirmed for at least six months showed no statistically significant differences between dimethyl fumarate and any comparator including

placebo. This analysis did not include Rebif 22 or natalizumab. Direction of effect favoured dimethyl fumarate except for the comparison with Betaferon (Table 1).

Table 1: Summary of results of MTC for ARR and EDSS progression based on Figure 21(P139) and Figure 28 (P145) in manufacturer's submission

	ARR: rate ratio (95% CI)	EDSS progression confirmed for at least three months: relative risk (95% CI)
Placebo	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]
Avonex	[redacted]	[redacted]
Betaferon	[redacted]	[redacted]
Rebif 22µg	[redacted]	[redacted]
Rebif 44µg	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]
Natalizumab	[redacted]	[redacted]
Teriflunomide 7 mg	[redacted]	[redacted]
Teriflunomide 14 mg	[redacted]	[redacted]

Table 2: Disability progression in dimethyl fumarate versus placebo groups (direct and indirect comparisons)

	Disability progression confirmed for at least three months: HR (95% CI)	Disability progression confirmed for at least six months: HR (95% CI)
DEFINE	0.62 (0.44 to 0.87)	0.77 (0.52 to 1.14)
CONFIRM	0.79 (0.52 to 1.19)	0.62 (0.37 to 1.03)
Pooled analysis	[redacted]	[redacted]
MTC	[redacted]	[redacted]

Statistically significant benefits on some quality of life measures and MRI outcomes (secondary or tertiary outcomes) were also documented in both trials.

Serious adverse events were uncommon. There were some types of adverse events which were more common in the dimethyl fumarate arms. These were flushing (and hot flushes), gastrointestinal (GI) events including abdominal pain, nausea and diarrhoea, and skin disorders (rash and pruritus). Increased incidences of these types of events were seen in both DEFINE and CONFIRM individually and in the pooled analyses of data from the two trials. Analysis by time-period of occurrence indicated that the majority of GI and flushing episodes occurred in the first three months of treatment and declined thereafter. This pattern of occurrence was seen in both the DEFINE and CONFIRM trials. There was no increased risk of opportunistic infection associated with dimethyl fumarate.

The ERG is aware of four cases of progressive multifocal leukoencephalopathy (PML) in patients treated with fumaric esters.

[redacted] The ERG's clinical advisor stated that, if PML were confirmed as an adverse event of dimethyl fumarate, guidance for discontinuation based on lymphocyte counts would be required and that monitoring of these parameters would therefore be required in clinical practice.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The submission included two RCTs which compared dimethyl fumarate with placebo, a pooled analysis of data from these trials and an MTC which included all the comparators defined as relevant by the NICE scope.

The ERG did not identify any relevant studies which were excluded from the submission. Both the phase III RCTs assessing dimethyl fumarate were at low risk of bias and were appropriately powered for placebo comparisons. Both trials had relevant primary outcomes and assessed secondary outcomes which were of key relevance to the decision problem. However the use of disability progression confirmed at three months as the primary measure of progression may be considered a non-ideal outcome assessment: assessment confirmed at six months is considered to be a more reliable indicator and was also presented.

Both DEFINE and CONFIRM had durations of 24 months. This was comparable to other trials in the field. However, this is a short period compared to the life-long course of multiple sclerosis where diagnosis is typically between the ages of 20 and 40 years and life expectancy is close to that of the general population. For patients with RRMS who are not treated with disease modifying therapy (DMT), progression to secondary progressive multiple sclerosis (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.

Neither trial was conducted primarily in the UK. Both DEFINE and CONFIRM were multicentre and multinational trials with worldwide recruitment. There were some differences in baseline characteristics between the trial populations and the UK clinical population of RRMS patients. The ERG's clinical advisor did not consider these differences likely to be clinically significant. The populations in all the trials (DEFINE, CONFIRM and the trials included in the MTC) are more closely representative of those patients who meet current Association of British Neurologists prescribing criteria for DMT than they are of the whole UK RRMS population.

The pooled analysis of the DEFINE and CONFIRM studies was conducted using standard methods and both fixed and random effects estimates were calculated. Statistical heterogeneity between the two trials was low and, although there were some differences between the two trial populations, the clinical characteristics of the studies were sufficiently similar to make calculation of a pooled estimate appropriate.

The MTC included 27 RCTs of eight treatments in patients with RRMS conducted over a period of 20 years. There was clinical heterogeneity between trials. However, the ERG did not consider that clinical heterogeneity was sufficient to make the network comparisons unreasonable. The manufacturer attempted to explore the impact of clinical variations using covariate analyses. The power of these analyses was limited by the small number of trials in many of the networks.

While a substantial number of trials were identified, many of the analyses contained only a minority of the total number of trials, based on availability of outcome data. Some of the networks were therefore sparsely populated and did not include all relevant comparators. In the analysis of three and six month confirmed disability progression at 24 months trials with shorter durations were excluded; the three month confirmed progression network did not include one of the comparator DMTs (Avonex) for this reason. Given the constraints of information available from trials assessing other DMTs, the MTC appeared complete.

Trials of fingolimod and of natalizumab which are licensed only for patients with rapidly evolving severe (RES) or highly active disease were included in the MTC. These therapies are recommended by NICE only for highly active and RES disease respectively. While these trials were conducted in broad RRMS populations they are relevant comparators only for these subgroups and estimates of effect for fingolimod and natalizumab versus dimethyl fumarate derived from the whole trial populations may not reflect the estimates in the indicated subgroups.

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

The manufacturer presented a *de novo* Markov model based upon a previously validated model evaluating treatments for MS. The model evaluated dimethyl fumarate compared to Rebif 22µg, Rebif 44µg, Avonex, Betaferon, glatiramer acetate, natalizumab and fingolimod in a RRMS population, over a time horizon of 30 years.

Disease progression was modelled using 21 health states, all of which represented different degrees of disease severity (through the progression in EDSS scores) whilst in RRMS and after conversion to SPMS and death. Whilst in RRMS, disability regression was possible. Both conversion to SPMS and regression within EDSS states once in SPMS were assumed irreversible. Treatments affect the health

of patients and cost to the health system through reduction in the annual relapse rate (ARR), the reduction in the annual risk of disability progression for a patient with RRMS, and through the occurrence of adverse events. Patients could discontinue the drug due to adverse events, moving to an EDSS state of 7 or higher, or through progression to SPMS. After withdrawing from any of the treatments modelled patients were assumed to receive no further treatment (i.e. placebo).

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. Outcomes were measured in terms of quality-adjusted life-years (QALYs) based on comparative effectiveness data and health-related quality-of-life (EQ-5D). Utility data were obtained from the pooled data from dimethyl fumarate trials, supplemented by the UK MS survey. Caregiver utilities were also considered. Resource use was derived from the UK MS survey and unit costs from relevant national sources were then applied.

Base case results were presented as deterministic pair-wise incremental cost-effectiveness results for dimethyl fumarate versus each of the comparators. In addition, a full incremental analysis was presented where drugs were compared to the next most expensive. This was undertaken for two scenarios, first where the list price was used for all drugs, and secondly where the manufacturer's proposed PAS price was used for dimethyl fumarate while using the list prices for all other drugs.

Deterministic sensitivity analyses generally showed that results were robust to those parameters tested. The manufacturer's base case analysis included the list prices for each drug. The ICER for dimethyl fumarate after conducting a full incremental cost-effectiveness analysis using the costs and QALYs derived from a probabilistic sensitivity analysis was £200,117 per QALY. The manufacturer also conducted a sensitivity analysis using the manufacturer's proposed PAS price and the list price for all other drugs.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The economic analysis presented by the manufacturer generally addressed the decision problem specified in NICE's scope. The structure of the model, although potentially limited due to its focus on EDSS, was sufficient to characterise the progressive nature of RRMS and adequately capture the majority of symptomatic and HRQoL aspects of the disease. Two of the comparators where not assessed within their licensed indications (fingolimod and natalizumab). No analysis was undertaken on the population subgroups for which fingolimod and natalizumab were licensed and recommended, so it is difficult to draw any conclusions regarding the comparative costs-effectiveness of dimethyl fumarate to either fingolimod or natalizumab in these subpopulations. The evidence used to populate

the model was derived from the two dimethyl fumarate trials, the UK MS Survey and the London Ontario dataset. The two latter sources have some limitations which introduce an element of uncertainty into the results, but they appear to represent the best available evidence.

The manufacturer reported results of several sensitivity analyses, including probabilistic analysis. However, these analyses are based on a model which utilises relative risks, rather than hazard ratios for progression outcomes and does not appear to have distributions assigned to all relevant parameters. Attempts to clarify this resulted in a model which appeared to be based on rate ratios, which the ERG believes to be more inappropriate than relative risks. Due to limitations in the availability of data and the lack of a full probabilistic model, there is some uncertainty in the ICER for dimethyl fumarate. Despite this, when the list prices are used for all drugs, the cost-effectiveness conclusion is robust to sensitivity analyses: dimethyl fumarate is not cost-effective given a cost-effectiveness threshold of £30,000 per QALY.

The manufacturer presented the results for an analysis where the manufacturer's proposed PAS price was used for dimethyl fumarate and the list prices for all other drugs. The ERG considers an analysis where reduced prices are used for all drugs where possible to be more appropriate.

The ERG considers probabilistic sensitivity analysis results to be appropriate rather than the deterministic results presented by the manufacturer.

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer

1.6.1 Strengths

The submission included evidence of effectiveness of dimethyl fumarate from two relevant good quality, moderate-size placebo-controlled trials (DEFINE and CONFIRM). Trial populations were broadly comparable to those patients who meet eligibility criteria for DMT in the UK.

These trials both showed a significant benefit of dimethyl fumarate versus placebo in reducing the ARR in patients with RRMS. Efficacy was confirmed by a pooled analysis of the placebo comparison from these two trials. Relevant comparators identified in the NICE scope were included in an MTC, which identified and included all the relevant trials; this demonstrated a benefit of dimethyl fumarate compared to all interferon therapies as well as glatiramer acetate, placebo and teriflunomide for ARR. The evidence from the two trials, the pooled analysis of these trials and the MTC consistently showed dimethyl fumarate to be effective in reducing relapse rates relative to other DMTs except fingolimod and natalizumab.

The outcome of three month disability progression showed a statistically significant benefit over placebo in one trial (DEFINE) and in the pooled analysis of DEFINE and CONFIRM; there was a non-significant benefit in the CONFIRM trial. The MTC also showed a statistically significant benefit for dimethyl fumarate over placebo. No statistically significant benefits were observed over active comparators but directions of effect favoured dimethyl fumarate. Although there was evidence of benefit, some uncertainty remains regarding the effect of dimethyl fumarate on disability progression, due to the limitations of this outcome measure.

Benefits in quality of life outcomes and MRI measures were identified in both DEFINE and CONFIRM, providing supportive evidence that dimethyl fumarate is associated with positive effects compared to placebo.

The economic analysis presented by the manufacturer generally addressed the decision problem specified in NICE's scope. The model structure is potentially limited due to its focus on the EDSS as this scale places greater emphasis on physical rather than cognitive changes and increments of one point represent ever greater changes in impairment as the scale increases. However, it was sufficient to characterise the progressive nature of MS and adequately capture the majority of symptomatic and HRQoL aspects of the disease and is consistent with previous submissions. In the model, regression to lower EDSS states was permitted for the RRMS population and modelled using data from the dimethyl fumarate trials. This is reasonable given that progression sustained for three months may not be permanent.

The model predictions in terms of mortality and the distribution of patients across EDSS states appeared reasonable compared to the two year trial data; although perhaps with a slightly higher proportion of patients in EDSS states 2 and 4 in the model output for dimethyl fumarate. This may slightly reduce progression over the long-run, which would favour dimethyl fumarate.

Trial population data were used where possible to inform natural history parameters, and the manufacturer appeared to use best available evidence where trial data were insufficient.

The economic model incorporated all the significant adverse events that occurred in the dimethyl fumarate trials. Some adverse events relevant to comparators may have been excluded but this is conservative with respect to dimethyl fumarate.

1.6.2 Weaknesses and areas of uncertainty

The duration of the key trials of dimethyl fumarate was two years, which is at the upper end of the range for MS trials (few trials in the MTC had a longer duration of follow-up), but is short in relation

to the duration of the disease. Two years is also substantially shorter than the period for which patients would be expected to be on disease modifying therapy (in RRMS patients may receive treatment with an EDSS score ≤ 7). There is therefore considerable uncertainty as to the long term efficacy and safety of dimethyl fumarate beyond two years. The economic model has a 30 year time horizon and the treatment effectiveness had to be extrapolated beyond the two year trial durations. The manufacturer incorporated a treatment waning effect and conducted sensitivity analyses around this, which was appropriate but uncertainty still remains.

Six month confirmed disability progression is considered a more robust measure than three month confirmed progression. Data for this measure were presented as sensitivity analyses for the trials of dimethyl fumarate, and showed less clear evidence of benefit than the three month confirmed progression. A statistically significant benefit compared to placebo was seen only in the pooled analysis of DEFINE and CONFIRM; individual trial confidence intervals included the possibility of no benefit. The MTC also showed no clear evidence of benefit with dimethyl fumarate over placebo for this measure.

Dimethyl fumarate has a CHMP positive opinion for use in all patients with a diagnosis of RRMS. Patients who are eligible for current DMTs (i.e. two clinically significant relapses in the previous two years for beta-interferons and glatiramer acetate) are a subset of those who meet diagnostic criteria for RRMS. [redacted] (as is the NICE scope) but the assessment of effectiveness which is contained in the submission relates more closely to patients who meet the criteria for current DMT, as the trials in the MTC had admission criteria requiring evidence of active relapsing disease in the baseline period. For example, in the trials of dimethyl fumarate, patients had to have had at least one relapse in the previous year to be eligible for treatment. Therefore the effectiveness of dimethyl fumarate in the whole clinical population of RRMS patients is uncertain.

The MTC addressed the decision problem of the relative effectiveness of dimethyl fumarate versus interferon-beta, glatiramer acetate, fingolimod and natalizumab in the population of RRMS patients. The manufacturer did not address the relative effectiveness of dimethyl fumarate versus fingolimod or natalizumab in the appropriate subgroups. The ERG accepts that this was due to lack of sufficient data on the efficacy of interventions in the RES and highly active subgroups. However the lack of analyses in these subgroups means that estimates of relative effect for relapse reduction and disability progression for these comparisons in the highly active and RES disease subgroups (respectively) are uncertain.

The evidence from the phase III RCTs indicated few serious adverse effects and a waning of the initially high levels of GI and flushing effects. However the ERG is aware of case reports of PML

occurring in patients treated with fumaric esters including dimethyl fumarate. It is unclear whether this risk is related to how these patients were managed.

Although an MTC analysis was conducted for adverse events, the results were not used in the economic model and this was not explained. Instead, incidence rates for each adverse event were calculated independently from the trials included in the systematic review of effectiveness evidence. However, the ERG conducted an analysis based on the results from the MTC and, although the ICER for dimethyl fumarate increased a little this was not significant.

There was uncertainty in the estimates of costs and utilities of EDSS states, relapses and adverse events. Many of these estimates were based on a population that was only partially comprised of RRMS patients or based on expert opinion. Furthermore, significantly different cost estimates for relapse and EDSS states were estimated in different submissions and publications based on the same population.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

As the ICER was so high for dimethyl fumarate when the list prices for all drugs were included, and the cost-effectiveness conclusions were robust to the manufacturer's sensitivity analyses at a threshold of £30,000 per QALY, the ERG conducted no further analyses assuming list prices for all drugs.

The manufacturer presented the results for an analysis where the manufacturer's proposed PAS price was used for dimethyl fumarate and the list prices for all other drugs. The ERG feel that this analysis is inappropriate and not in line with NICE methods guidance. The ERG considers an analysis where reduced prices are used for all drugs where possible to be more appropriate. The ERG conducted an analysis assuming the risk sharing scheme prices for Rebif 22µg, Rebif 44µg, Avonex, Betaferon and glatiramer acetate published in a Circular by the Department of Health in 2002, an estimate 35% price reduction for fingolimod (the actual price reduction is not publically available), the list price for natalilzumab and the manufacturer's proposed PAS price for diemthyl fumarate. The deterministic ICER for dimethyl fumarate was £36,511 per QALY. The more appropriate probabilistic sensitivity analysis result was £49,687 per QALY.

All further analyses undertaken by the ERG used the discounted prices for all drugs. The deterministic ICERs were calculated for each of these analyses rather than the results from probabilistic sensitivity analyses because of the computation time required to run the probabilistic sensitivity analyses. The change in the deterministic ICER from the ERG base case using discounted prices for all drugs where appropriate should be related to the base case result of £49,687 per QALY when interpreting the importance of the change.

The analyses are listed below with the ICER results for dimethyl fumarate. These results are deterministic. The ERG base case deterministic ICER for dimethyl fumarate is £36,511.

- Alternative treatment monitoring resource assumptions: ICER ranged from: £37,477 to £43,874;
- Discontinuation rate after two years is 50% or 0% of the trial duration discontinuation rate for dimethyl and the comparator: ICER ranged from: £40,633 to £48,436;
- Using the 95% lower and upper limits of the confidence interval for relative discontinuation risks for dimethyl fumarate versus glatiramer acetate: ICER ranged from: £31,367 to £40,546;
- Transition rates to SPMS for each EDSS state increased or decreased by 50%: ICER ranged from: £34,345 to £39,568;
- Alternative utility estimates for EDSS states using other publications: ICER ranged from: £34,427 to £37,952;
- Alternative cost estimates for EDSS states using other publications: ICER ranged from: £32,157 to £39,248;
- Natural history relapse rates from MS survey: ICER was £38,356;
- Alternative relapse cost estimates from other publications: ICER ranged from: £35,116 to £38,923;
- No adverse events assumed: ICER was £37,818;
- Adverse events derived from MTC: ICER was £37,176;
- Alternative utility estimates for flu-like symptoms and influenza: ICER was £36,504.

2 Background

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

The manufacturer's description of the multiple sclerosis (MS) appeared appropriate and relevant. It correctly characterised the different forms of the disease and the disease course which leads to the development of secondary progressive MS (SPMS) after a period of years with relapsing remitting MS (RRMS).

The number of adults with RRMS in England and Wales was estimated at 51,749 based on the 2011 census data which recorded the number of adults in these countries as 43,486,200.[1, 2] This represents the upper boundary of the MS prevalence estimate (of 74-140 per 100,000); and the proportion of MS patients who have RRMS at diagnosis (85%).[2] Independent verification by the ERG with an alternative source indicated that the use of the upper bound of this prevalence estimate was reasonable.[3] However, applying an incidence of 85% of patients with RRMS at diagnosis to estimate the prevalence of the relapsing form of the disease in the population may have overestimated the prevalent population. The natural history of the disease indicates that approximately 50% of patients initially diagnosed with RRMS will convert to SPMS within 10 years. The proportion of adults with MS who have a relapsing remitting form of the disease at a given time is therefore lower. The ERG's clinical advisor indicated that the figure is likely to be closer to 60%. Applying this to the estimate of MS prevalence would generate a figure of 36,528 adults with RRMS. The comparable figure from the SWIMS study, which is a prospective longitudinal study of people with MS in Devon and Cornwall, gave an estimate of 36% but there was a substantial proportion of patients who did not classify their disease (21%).[4]

2.2 Critique of manufacturer's overview of current service provision

The manufacturer accurately summarised the lack of a well-defined clinical pathway for the treatment of patients with RRMS and appropriately characterised the fact that the majority of patients initiate treatment with interferon beta-1a/b (Avonex, Rebif 22, Rebif 44, Betaferon) or glatiramer acetate, and that a minority of patients with rapidly evolving severe (RES) disease commence with nataluzimab.[5, 6] Fingolimod was appropriately identified as second-line treatment, recommended for highly active disease where relapses have not been controlled by treatment with beta-interferon.[7] The fluidity of treatment selection was also described, accurately explaining the potential for patients to switch between forms of interferon-beta and/or glatiramer acetate or to escalate to therapy with natalizumab or fingolimod if additional criteria were met. The role of patient preference and adverse effect profiles was also referred to. Although the background clearly describes the criteria for prescribing of fingolimod and natalizumab, the criteria for prescription of interferon beta-1a/b and glatiramer acetate were not explicitly stated. These disease modifying therapies (DMTs) are prescribed according to the

Association for British Neurologist Guidelines.[8] The guideline states that eligible patients for treatment are normally ambulant adults (maximum EDSS score of 6.5) with active relapsing disease defined as two clinically significant relapses in the previous two years.[8]

In the submission the company requested that dimethyl fumarate is accepted for use within its full licensed indication: it currently has an EMA Committee for Medicinal Products for Human Use (CHMP) positive opinion for its use in the treatment of adults with RRMS. The patients they identify as eligible for dimentyl fumarate are: treatment naïve patients (including newly diagnosed and those who have not received a DMT); those who have discontinued a previous treatment due to lack of efficacy or tolerability; and patients who are sub-optimally treated or dissatisfied with their current treatment. They correctly state that there is currently no oral therapy available for first-line use in RRMS and identify dimeythl fumarate as a therapy that can be used first-line, as a non-injectable alternative for patients currently on interferon beta or glatiramer acetate, and for patients unwilling to self-inject. They also propose dimethyl fumarate as an alternative second-line therapy to natalizumab and fingolimod.

In their assessment of implementation the manufacturer's submission stated (P276) that it was assumed that 100% of the eligible population would receive treatment with interferon beta or glatiramer acetate. The implication was that 100% of adults with RRMS (the licensed indication for dimethyl fumarate) would be currently receiving this therapy. As described above, under current guidelines 100% of the prevalent population are not eligible for treatment with interferon beta or glatiramer acetate: some patients with RRMS have relapses too infrequently to meet prescribing criteria for currently available DMTs. There are therefore a substantial proportion of patients who would be eligible for dimethyl fumarate under the current CHMP draft opinion who are currently receiving no DMT. These patients may be in receipt of best supportive care, which is not included as a comparator in the NICE scope or in the manufacturer's submission.

The manufacturer also assumed in their estimation of patients eligible for treatment that no patients currently on fingolimod or natalizumab would switch to dimethyl fumarate; it is not clear to the ERG if this is a reasonable assumption. The manufacturer also makes the assumption in the economic model, based on current guidelines, that treatment will cease when a patient's EDSS score reaches ≥7.0. The Association of British Neurologists suggest treatment cessation at EDSS 7.0 with the development of secondary progressive MS.[8] The ERG's clinical advisor regarded the assumption about treatment cessation as reasonable. However, it is unclear if alternative treatments would be considered rather than all treatment ceasing.

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).

The clinical evidence presented in the submission consists of trials conducted in a population of adult RRMS patients, both in the studies assessing dimethyl fumarate (DEFINE and CONFIRM) and those in the MTC which assess comparator DMTs. Clinical advice suggests that the populations of the two placebo-controlled RCTs of dimethyl fumarate were broadly comparable with the population of RRMS patients seen in UK clinical practice in age profile, disease duration and ratio of female to male patients. Patients who receive interferon beta or glatiramer acetate in the UK are in the NHS risk sharing scheme (RSS). They therefore meet the ABN prescribing guidelines and have higher relapse rates than the prevalent population.[9] Compared to patients in the RSS cohort, patients in the two dimethyl fumarate trials had slightly less disability as measured by the EDSS: in the RSS the mean EDSS score was 3.1(SD 1.5) for RRMS patients whereas in the trials the scores were slightly lower (CONFIRM mean 2.4, SD 1.17; DEFINE mean 2.6, SD 1.24). The baseline annualised relapse rate (ARR) in the trials (1.3 in DEFINE and 1.4 in CONFIRM) was higher than the mean for the UK clinical population in the view of the ERG's clinical advisor. The SWIMS study did not report mean ARR for RRMS patients but the majority of all MS patients reported 0 or 1 relapse in the previous year.[4] However, the trial baseline ARR reflects the inclusion criterion that patients needed to have had ≥1 relapse in the previous 12 months. This criterion is different to UK prescribing guidelines for current DMT, which states that patients should have had ≥2 relapses in the preceding two years.[8]

The NICE scope identified the following subgroups as being of interest:

i) Patients with RRMS whose disease has inadequately responded to treatment with DMT
ii) Patients with RRMS whose disease is intolerant to treatment with DMT
iii) Patients with highly active RRMS
iv) Patients with RES RRMS

The scope stated that these groups should be considered if evidence allowed. The manufacturer's submission did not consider these subgroups in either the assessment of clinical effectiveness or the economic model. The rationale stated for this was that dimethyl fumarate is indicated for the whole adult RRMS population.

[redacted]

It could be contended that since only patients who meet the criteria for (iii) and (iv) are eligible for treatment with fingolimod and natalizumab respectively, a subgroup analysis would have been appropriate for these specific comparisons. The key trials of natalizumab (AFFIRM) and fingolimod (FREEDOMS and TRANSFORMS) were conducted in general RRMS populations and are therefore relevant for inclusion in the main analysis.[10-12] Nevertheless the usefulness of the results from these analyses is open to question: the MTC effect estimates for dimethyl fumarate versus natalizumab and fingolimod from the whole trial populations do not represent those for the relevant subgroups as defined in the NICE scope nor do they represent the estimates used in informing the licenses or NICE recommendations; these were informed by subgroup analysis.[6, 7, 13] The uncertainty around this is discussed in section 4.3.1.6. Therefore the decision problems of dimethyl fumarate versus fingolimod in patients with highly active disease and dimethyl fumarate versus natalizumab in patients with RES disease were not addressed by the submission.

The prespecified subgroup analyses conducted and presented by the manufacturer for the dimethyl fumarate trials included comparison of patients previously treated with DMT or alternative therapies and those who were treatment naïve.

3.2 Intervention

Dimethyl fumarate does not currently have a UK marketing authorisation. The intervention described in the submission is that of oral dimethyl fumarate. Although the NICE scope did not specify the dose; the dose in the submission was 240mg, twice daily (BID) which is in accordance with the provisional CHMP approval. The indication in the CHMP approval is that of adults with RRMS which matches that stated in the scope.

3.3 Comparators

The NICE scope defined the relevant comparators as being beta-interferon, glatiramer acetate, natalizumab and fingolimod. 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). Despite not being 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.[13] Both natalizumab and fingolimod have been licensed and evaluated by NICE subsequent to the NHS RRS, so are not part of the scheme; however in the case of fingolimod a PAS in in operation.

Beta-interferon and glatiramer acetate are both licensed and recommended for first-line treatments of the RRMS population. However, natalizumab and fingolimod are defined as relevant only for subgroups of the RRMS population who meet the criteria for rapidly evolving severe RRMS

(natalizumab) or highly active RRMS, in accordance with their licenses and NICE guidance for their use (TA127 and TA254 respectively).[6, 7]

The comparators described in the submission match those identified in the scope. The MTC included trials of both interferon beta-1a and interferon beta-1b in their relevant forms at clinically relevant/approved doses (see Table 3). Two forms of interferon beta-1a (Avonex and Rebif) were included and both licensed doses of Rebif (22µg and 44µg) were assessed; only one form of interferon beta-1b (Betaferon) was included. As discussed in section 3.1, comparisons with fingolimod and natalizumab were not restricted to their licensed or recommended indications as set out in the NICE scope but were included as comparators for all RRMS patients. Estimates of effect derived from these comparisons are not applicable to the indicated populations.

The submission did not define the alternative novel oral agent teriflunomide as a comparator, but trials of teriflunomide versus placebo were included in the MTC and effectiveness results for dimethyl fumarate versus teriflunomide were presented. Trials of the novel oral agent alemtuzumab were specifically excluded from the MTC searches (see section 4.1.2).

Table 3: Comparators identified in the systematic review inclusion criteria. Based on Table 5 (P37) in manufacturer's submission

Comparator	(Formulation) Dose
Interferon beta-1a	Avonex 30µg weekly intramuscular (i.m.)
Interferon beta-1a	Rebif 44µg subcutaneous (s.c.)
Interferon beta-1a	Rebif 22 µg 44µg s.c.
Interferon beta-1b	Betaferon 250µg s.c.
Glatiramer acetate	20mg daily s.c.
Natalizumab	300mg monthly intravenous (i.v.)
Fingolimod	0.5mg daily oral (p.o.)
Teriflunomide	Not yet licensed by EMA*

* Both 14mg (FDA licensed dose) and 7mg were included in review.

3.4 Outcomes

The NICE scope defined the relevant outcomes as relapse rate, severity of relapse, disability, symptoms of MS, freedom from disease activity, mortality, adverse effects of treatment and health-related quality of life (HRQoL).

The submission addressed all these outcomes with the exception of severity of relapse. The rationale for excluding this was that the trials of dimethyl fumarate did not evaluate this outcome. The rate of relapses requiring IV steroid treatment was assessed as a tertiary outcome in the two trials but was limited in its usefulness by the fact that IV steroids were the only protocol-allowed therapy (and

therefore relapses which may have been treated with oral steroids in clinical practice were included in this measure). Additional outcomes assessed in the clinical evidence submission (though not the economic model) included MRI outcomes such as T1, T2 and Gd+ lesions; the rationale for their inclusion was that they may indicate disease activity in the absence of relapses/disability progression.

3.4.1 Relapses

ARR was the primary outcome of the CONFIRM study and a secondary outcome in the DEFINE study. The EMA stated that this is an acceptable parameter to assess relapses in a guideline on clinical investigation of medicinal products for the treatment of MS that has been recently out for public consultation.[14] The primary outcome in the DEFINE study was the proportion of patients relapsing at 24 months. The EMA guidance further states that efficacy should be demonstrated over a period of at least two years, that relapses be clearly defined and that corticosteroid treatment for relapses should be carefully standardised. The ERG considers that this was the case in both the DEFINE and CONFIRM trials. Additional outcomes relating to the proportion of patients free of relapse/with relapse were also reported for both one and two years.

3.4.2 Sustained disability progression

The main measure for sustained disability progression in both the trials and the economic model was increase in EDSS score. This is a 20 point scale ranging from 0 (normal neurological examination) to 10 (death) in 0.5 increments. The EDSS is the most widely used measure of disability and its progression in MS and its use is recommended by the EMA.[14] However, the EDSS has well-documented limitations. Although it assesses seven functional systems (pyramidal, cerebellar, brainstem, sensory, bowel/bladder, visual, cerebral and other) which contribute to the calculation of scores, it underestimates the impact of cognitive changes and in the higher score levels it is driven largely by mobility decrements. It is notable that a one point change from a low score does not represent the same change in impairment as a one-point change from a higher score. The ERG's clinical advisor made the point that at the higher levels of the EDSS an increase of 0.5 can be highly clinically significant, whereas a 0.5 increase at the lower levels might be less important clinically.

EMA guidance stated that time to EDSS progression or the proportion of individuals showing progression at a pre-specified time are both acceptable parameters for assessing disability progression.[14] It also recommends an interval of at least six months between two assessments on the EDSS to establish whether any deterioration on the scale is sustained (i.e. represents permanent disability progression).

The prespecified secondary outcome in both DEFINE and CONFIRM (and the main driver of the model) was sustained disability defined as an increase of ≥1.0 point in the EDSS score sustained for at least three months from a baseline score of ≥1.0 (or ≥1.5 from a baseline score of 0). The definition used of an increase in one point also matches the EMA recommendation of one point being of relevant magnitude when the baseline EDSS score is ≤5.5. This measure has been used in previous appraisals (TA254)[7] but is not consistent with the EMA guidance and the advice of the ERG's clinical advisor that six month confirmed progression may be a more reliable measure of disability progression, as at three months there is still the possibility of recovery to a lower EDSS score. Data were also presented for analyses of progression sustained for at least six months which were undertaken as prespecified sensitivity analyses of this outcome.

EMA draft guidance also recommends the use of alternative additional measures of disability progression. Alternative measures presented for both the DEFINE and CONFIRM trials were differences from baseline in the Multiple Sclerosis Functional Composite (MSFC) score and its individual measures (timed 25-foot walk test, the 9-hole peg test and the paced audio serial addition test). Progression of cognitive deficit and change in visual function were also assessed in both trials.

3.4.3 Other outcomes

Quality of life was assessed using a global well-being visual analogue scale (VAS) with 0 as poor and 100 as excellent, the SF-36, the EQ-5D and the EQ-VAS. Results were reported separately for individual elements of the SF-36 and summary component scores. MRI outcomes reported were the number of new T1 or T2 lesions, the volume of T1 or T2 lesions, number of gadolinium enhancing (Gd+) lesions, brain atrophy and conversion of Gd+ to T1 lesions. These measures include those identified as relevant by the EMA.

3.5 Other relevant factors

A patient access scheme application was submitted along with the documentation. This is awaiting Department of Health approval.

4 Clinical Effectiveness

This section contains a critique of the methods of the manufacturer's reviews of clinical effectiveness data, followed by a description and critique of the trials included in the review, including a summary of their quality and results and the results of any synthesis of studies.

A systematic review of RCTs of dimethyl fumarate and the comparators stated in the NICE scope, with the addition of teriflunomde, was undertaken. A systematic review of non-RCTs was also

presented, although this was limited to studies which included dimethyl fumarate as a comparator. The main components to the ERG critique of the clinical effectiveness evidence are:

1) A critique of the systematic review methods;
2) A brief description of the key effectiveness results from two dimethyl fumarate placebo-controlled RCTs (CONFIRM and DEFINE) and a critique of the evidence;
3) A brief summary of the results from the manufacturer's meta-analysis of the CONFIRM and DEFINE trials and a critique of the evidence;
4) A brief summary and critique of the mixed treatment comparison (MTC) undertaken of dimethyl fumarate and other DMT's and placebo, and
5) A summary of tolerability and safety.

4.1 Critique of the methods of review(s)

4.1.1 Searches

The manufacturer's submission adequately described the search strategies used to identify relevant studies relating to the use of dimethyl fumarate for RRMS. Full details of the strategies used in each section were reported in the appendices of the submission or in the clarifications provided by the manufacturer in response to queries raised by the ERG.

Overall, the search strategies employed for each of the clinical effectiveness sections of the submission were appropriate and well documented. There were some weaknesses in the strategies, however it is unlikely that any of these would lead to relevant studies being missed by the searches. A detailed commentary on the individual searches is provided in the Appendix (section 10.1.1).

4.1.2 Inclusion criteria

Clear inclusion criteria were stated for the systematic review of trials of dimethyl fumarate and relevant comparators. These are briefly summarised in Table 4 below.

The study selection process was carried out in duplicate by two independent reviewers at both the initial stage of title and abstract and with full text studies; disagreements were resolved by a third reviewer. This was an appropriate method of study selection.

Table 4: Inclusion criteria for systematic review of trials of dimethyl fumarate and specified comparators. Based on Table 5 (P37) in manufacturer's submission

Population	Adults aged ≥18 years with RRMS (≥80% trial population)
Intervention	Licensed dose of Interferon beta-1a Interferon beta-1b Glatiramer acetate

	Dimethyl fumarate* Fingolimod Natalizumab Teriflunomide
Comparator	Any other included intervention also at licensed dose Placebo Best supportive care
Study design	RCTsand non RCTs for dimethyl fumarate only
Other	Studies with mixed populations (disease/age) required to report subgroup data for population of interest. Published before October/November 2012 Published in English

*Non RCTs also eligible for dimethyl fumarate

For population, intervention, comparator and dose, uniform inclusion criteria were used. However for study design, the inclusion criterion differed between dimethyl fumarate and the defined comparators, with non-RCTs eligible only if they assessed dimethyl fumarate.

The inclusion criteria were appropriate to the purpose of the review. The lack of requirement for blinding as a criterion was appropriate to ensuring completeness of the data set, particularly given the fact that the majority of the comparator DMTs are delivered by injection, and blinding is often considered inappropriate in these contexts.

The ERG asked the manufacturer to comment on the specific exclusion of alemtuzumab and laquinimod as comparators whereas the out of scope teriflunomide was included. The manufacturer's response stated that they included only licensed interventions and their approved doses for the treatment of RRMS; they stated that teriflunomide was specifically included because it received FDA approval prior to the review dates. The ERG notes that the 7 mg dose of teriflunomide was included together with the FDA licensed 14 mg dose; the impact of this on results is likely to be insignificant.

The use of a language restriction, with only studies reported in English has the potential to lead to selection bias (as well as the more general omission of relevant studies) but is listed as being due to NICE preference. The ERG was unable to verify the source of this preference.

In order to verify the application of the inclusion criteria to the identified studies, the ERG requested that the manufacturer provide the list of studies excluded at full text screening, together with reasons for their exclusion. This list of studies excluded at the final stage was supplied and checked by the ERG: it did not contain any studies which should have been included. See section 4.2.5 for further discussion of studies excluded from the review of dimethyl fumarate and section 4.3.3 for studies excluded from the MTC.

It was unclear from the submission whether any relevant non-RCTs of dimethyl fumarate were included (two were noted as being identified): the ERG requested clarification on either a) details of

the included studies or b) justification for their exclusion. The manufacturer clarified that the two non-RCTs initially identified were subsequently assessed as not being relevant to the submission. Having assessed these studies, the ERG agreed that this decision was correct.

4.1.3 Critique of data extraction

The methods used for data extraction involved reasonable measures to reduce reviewer error or bias, with data entered by one reviewer checked by a second and disagreements resolved through discussion. The ERG checked main outcome data against published trial reports and, where appropriate, the clinical study reports supplied by the manufacturer. These were accurately reported with one exception.

4.1.4 Quality assessment

The trials were assessed for quality in the manufacturer's submission using criteria which broadly reflect those of the Cochrane risk of bias tool.[15] Items relating to blinding of patients, personnel and outcome assessors were grouped as one question although substantiation of the answers referred to all three. It was unclear whether the assessment had been conducted in duplicate. The submission assessed both the DEFINE and CONFIRM trials as meeting all these quality criteria (manufacturer's submission, section 6.4.3, P68) and provided substantiation for these assessments (manufacturer's submission, section 10.3, P297).

The ERG replicated the quality assessments based on the totality of information available, including the published papers and protocols for the DEFINE and CONFIRM studies.[16, 17] The ERG's quality assessment, using the manufacturer's criteria is shown in Table 5.

Table 5: ERG assessment of dimethyl fumarate trials using manufacturer's criteria.

	DEFINE	CONFIRM
Was randomisation carried out appropriately?	Yes Centralised interactive voice response system Stratified by site	Yes Centralised interactive voice response system Stratified by site
Was the concealment of treatment allocation adequate?	Yes Centralised interactive voice response used six digit code to allocate treatment	Yes Centralised interactive voice response used six digit code to allocate treatment
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes There was good comparability of groups across key variables including relapses in year prior to randomisation, baseline EDSS scores and previous treatment with DMT.	Yes There was good comparability of groups across key variables including relapses in year prior to randomisation, baseline EDSS scores and previous treatment with DMT.
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes Patients in placebo and DF groups instructed not to take medication for 4 hours before study visits to prevent flushing reactions resulting in unblinding. This was a reasonable approach, though it may not have prevented unblinding of patients.	Yes for DF versus placebo comparison No for glatiramer acetate versus placebo or glatiramer acetate versus DF. Patients in placebo and DF groups instructed not to take medication for 4 hours before study visits to prevent

	Separate examining and treating neurologists at each site. INEC used blinded clinical records without MRI data	flushing reactions resulting in unblinding. This was a reasonable approach, though it may not have prevented unblinding of patients Separate examining and treating neurologists at each site. INEC used blinded clinical records without MRI data
Were there any unexpected imbalances in drop-outs between the groups	No Discontinuation rates for adverse events and relapses differed between the groups but there were no unexpected differences and overall discontinuation did not significantly differ*	No Discontinuation rates for adverse events and relapses differed between the groups but there were no unexpected differences and overall discontinuation did not significantly differ*
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No The full CSR was supplied by the manufacturer and all outcomes in the published trial protocol are included.	No The full CSR was supplied by the manufacturer and all outcomes in the published trial protocol are included.
Did the analysis include an intention-to-treat analysis? If so was this appropriate and were appropriate methods used to account for missing data?	Yes The primary efficacy analysis was all randomised patients who received ≥1 dose of medication. This was 1234 patients compared to 1237 randomised. Per protocol, MRI cohort and safety analyses were also documented	Yes The primary efficacy analysis was all randomised patients who received ≥1 dose of medication. This was 1417 patients compared to 1430 randomised. Per protocol, MRI cohort and safety analyses were also documented†

*Pooled analysis presented in submission showing statistically significantly higher discontinuation for AE in DF groups but lower rates for any cause.
†Withdrawals/treatment switchers in CONFIRM were censored and therefore excluded from primary analysis of confirmed disability progression (data were not censored in DEFINE) but sensitivity analyses including them were conducted.

The ERG was in agreement with the manufacturer's overall assessment of study quality, although they did note a small difference in the completion of the checklist. The manufacturer's submission scored the blinding item with "YES", although noting that the patients enrolled in the glatiramer acetate arm were unblinded; the examining neurologist was blinded to treatment for all patients, including those receiving glatiramer acetate. In addition the ERG noted that the trial was not powered to assess the comparison between dimethyl fumarate and glatiramer acetate and that the manufacturer's assessment was accurate with respect to the placebo comparison.

[redacted]. However, sensitivity analyses using site-assessed objective relapses were also conducted and these did not significantly alter results.

Power does not form part of the risk of bias assessment but is a relevant factor. Both trials were adequately powered to assess the placebo comparison for dimethyl fumarate 240 mg BID versus placebo for the primary outcome. In the case of DEFINE the study was designed to have at least 90% power to detect a 30% reduction in the proportion of patients relapsed at two years between the dimethyl fumarate and placebo groups. CONFIRM was designed to have 84% power to detect a 25% reduction in the ARR between the dimethyl fumarate and placebo groups over two years.

4.1.5 Evidence synthesis

The manufacturer also submitted a pooled analysis of DEFINE and CONFIRM dimethyl fumarate versus placebo comparisons using both fixed (Mantel-Haenszel) and random (Dersimonian and Laird) effects for the following efficacy outcomes: ARR, ARR for steroid-treated relapses, proportion of patients with relapse at 12 months, proportion of patients with relapse at 24 months, proportion of patients remaining relapse free at 12 months, proportion of patients remaining relapse free at 24 months; change in EDSS score at 24 months; disability progression sustained for 3 months at 24 months, and disability progression sustained for 6 months at 24 months. Rate ratios, relative risks or weighted mean differences were presented. In the original submission the manufacturer undertook the meta-analysis for disability progression using relative risk rather than hazard ratio which was used for the individual trials and is a more appropriate analysis for confirmed disability progression. In the clarifications submitted by the manufacturer they re-ran the meta-analysis using the hazard ratio for disability progression and the ERG have used the revised analysis. The appropriateness of the pooled analysis is considered in section 4.2.3.

Meta-analysis results were also presented for the safety and tolerability outcomes assessed in the two trials. These were presented in place of results for the individual RCTs: the decision to provide only pooled safety data in the submission appeared appropriate.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Trials included in the submisison

Two phase three randomised controlled trials (RCTs) which compared dimethyl fumarate (Tecfidera™) in its approved dose of 240mg BID with placebo in patients with RRMS (the DEFINE and CONFIRM trials) were included in the systematic review undertaken by the manufacturer. The CONFIRM study also used an active comparator, glatiramer acetate at the approved dose of 20 mg given once daily by subcutaneous (sc) injection. As stated by the manufacturer in the submission glatiramer acetate was included as a reference group only. The study was not designed to test the superiority or non-inferiority of dimethyl fumarate versus glatiramer acetate.[18] Based on the statistical analysis plan in the study protocol it was a tertiary objective to compare the benefit-risk profile of dimethyl fumarate versus placebo with glatiramer acetate versus placebo.[16] Any comparisons of dimethyl fumarate versus glatiramer acetate were post-hoc.

The duration of both studies was 96 weeks. Both trials also included an arm treated with a higher dose (240 mg TID) of dimethyl fumarate. The primary outcome in DEFINE was the proportion of patients experiencing relapse by two years; the primary outcome in CONFIRM was annualised relapse rate (ARR). Both trials reported sustained EDSS progression as a secondary outcome. Key baseline characteristics of these trials are shown in Table 6.

Table 6: Population characteristics of the DEFINE and CONFIRM trials. Adapted from Table 10 (P54) and Table 30 (P122) in manufacturer's submission

	DEFINE			CONFIRM			
Trial Arm	Dimethyl fumarate 240 mg BID	Dimethyl fumarate 240 mg TID	Placebo	Dimethyl fumarate 240 mg BID	Dimethyl fumarate 240 mg TID	Placebo	Glatiramer acetate 20 mg OD
N (ITT)	411	416	410	362	345	363	360
N (modified ITT)	410	416	408	359	345	362	350
Mean age: years	38.1	38.8	38.5	37.8	37.8	36.9	36.7
Females: N (%)	136 (77)	140 (76)	141 (78)	118 (70)	121 (71)	123 (70)	116 (69)
Mean EDSS	2.40	2.36	2.48	2.56	2.52	2.59	2.57
Disease duration: mean	5.6	5.1	5.8	4.9	4.6	4.8	4.4
History of DMT: N (%) Interferon beta-1a Interferon beta-1b Glatiramer acetate Natalizumab	 114 (28) 57 (14) 52 (13) 8 (2)	 111 (27) 60 (14) 60 (14) 18 (4)	 106 (26) 55 (13) 75 (18) 8 (2)	 66 (18) 42 (12) 1 (<1) 2 (<1)	 70 (20) 39 (11) 3 (<1) 6 (2)	 80 (18) 43 (12) 1 (<1) 6 (2)	 76 (22) 33 (9) 1 (<1) 2 (<1)
Region Region 1 Region 2 Region 3							
Mean relapses in year prior to randomisation	1.3	1.3	1.3	1.3	1.4	1.4	1.4

4.2.2 Population and relevance to UK

Both the DEFINE and CONFIRM studies were multinational studies conducted across 28 countries in North and Central America, Asia, Eastern and Western Europe, the Middle East and Australasia, using 198 (DEFINE) or 200 (CONFIRM) centres. The DEFINE study included only 29 patients from seven UK centres; the CONFIRM study did not enrol any UK patients. The majority of patients in both studies were White (79% and 84% in DEFINE and CONFIRM respectively), with minorities of Asian and Black participants as well as a minority whose race was not recorded. Breakdowns by region of recruitment were provided for the following groupings:

[redacted] The ERG's clinical advisor indicated that patients from Asia, where the incidence of MS is considerably lower,[3] may have a different disease profile to those from Western Europe or North America. However, the regional groupings as presented mean that it would be difficult to identify any differential estimate of effect in such patients.

[redacted]

4.2.2.1 Age and sex profile of patients.

The trial populations reflect those of the UK clinical population in having a majority of female patients (74% in DEFINE and 70% in CONFIRM).[4, 9] Age profiles of patients in the two trials also reflect the disease course in which most patients are diagnosed between the ages of 20 and 40: a majority of patients in both trials were aged 30 to 39 or 40 to 55 years, with almost none aged over 55. Mean ages were 38.5 years for DEFINE and 37.3 in CONFIRM.

4.2.2.2 Previous treatment

There were baseline differences between the two trials in the proportion of patients who had previously taken a DMT. 55% of patients in the DEFINE study had had prior therapy for MS, of which 41% had received a DMT. In the CONFIRM trial 40% had received prior therapy, with the number having had DMT being 29%. The proportion who had received DMT in DEFINE may be higher than would be the case for the UK clinical population; in a longitudinal study of MS in one English region 31% of the RRMS patients studies received current or prior DMT with 23% of patients reporting currently receiving therapy.[4]

The most common prior treatment in both trials was interferon beta-1a (not specified whether Avonex or Rebif or dose for Rebif) with smaller numbers receiving interferon beta-1b. In the DEFINE trial an approximately equal number of patients had received glatiramer acetate as had had interferon beta-1b but in the CONFIRM study almost no patients had received glatiramer acetate because it was a prohibited study medication. Small numbers of patients in DEFINE had received natalizumab, these figures were lower for the CONFIRM study. It appeared that most patients had received only one DMT. This may not be representative of the current UK population of RRMS patients, where it is not uncommon for patients to switch DMT due to adverse effects or lack of clinical efficacy (acknowledged by the manufacturer in their characterisation of the current care pathway (manufacturer's submission, section 2.5). Subgroup analysis (see section 4.2.4.4 below) of treatment naïve and treatment experienced patients in each trial found broadly consistent results for the two groups in CONFIRM although in DEFINE there was some evidence of a trend towards dimethyl fumarate having greater efficacy in the treatment naïve group. It should be noted that these subgroup analyses grouped patients whose prior MS treatment was and was not a DMT together in the prior treatment subgroups.

4.2.2.3 Baseline disease characteristics

Mean EDSS scores (between 2.36 and 2.59 in treatments groups) were slightly lower than the population participating in the NHS RSS (mean EDSS for RRMS patients: 3.1).[9]

The mean baseline relapse rate was 1.3 in the DEFINE and 1.4 in the CONFIRM trial This reflected the inclusion criterion in both trials that patients had ≥1 relapse in the year prior to randomisation The number of relapses in the year prior to randomisation in both trials was comparable to that of patients in the RSS who had a mean of 2.9 (SD 1.2) relapses in the previous two years but somewhat higher than the whole UK RRMS population; the ERG's clinical advisor suggested that a typical ARR in clinical practice would be in the region of 0.8. In the context of the ABN prescribing guidelines for currently available DMTs, of at least two clinically significant relapses in the previous two years, the baseline mean relapse rate was considered reasonable by the ERG's clinical advisor. The placebo groups ARRs, adjusted for baseline characteristics at 24 months, were 0.36 in DEFINE; and 0.40 in CONFIRM. Given the inclusion criteria of the trials which require ≥1 documented relapse in the previous year, it is possible that patients were recruited at a point with atypically (for them) high relapse rates. It is also possible that, as ARR is known to be negatively correlated with disease duration, that it would be expected to drop naturally over the course of a two year trial. The impact of strict criteria for confirmed relapses in an RCT may also play a role. However, an adjusted ARR of 0.4 for a placebo group is representative of some of the more recent trials in the MTC.[19, 20]

4.2.3 Statistical analyses

Analyses presented in the submission for primary outcomes and ARR (in the case of DEFINE) and for confirmed disability progression for each trial were checked against the planned analyses as outlined in the statistical analysis plan in the trial protocols.[16, 17] No discrepancies were identified between the primary analysis and sensitivity analysis specified in the protocols for DEFINE and CONFIRM and what was reported in the submission. Subgroup and sensitivity analyses presented in the submission were identified as pre-planned or post-hoc.

Pooled analysis was conducted using both fixed and random effects (see section 3.1.5), which was appropriate. There were some differences between DEFINE and CONFIRM in population characteristics (see section 3.2.2.) These primarily related to the regions patients were recruited from and to prior treatment.

[redacted]. DEFINE had higher proportions of patients with any previous therapy and with previous DMT. Additionally CONFIRM did not include patients with a history of glatiramer acetate therapy.

[redacted]. Likewise there were some methodological differences between the trials such as rules for switching medication and the handling of withdrawals for EDSS analyses which is of particular relevance to the pooling of the two studies for 3 and 6 month confirmed disability progression. This is discussed in more detail in 4.2.5.

[redacted]

There was almost no difference between the estimates of effect between the random and fixed effect models. Statistical heterogeneity was assessed using the I^2 statistic. $I^2 = 0\%$ for all analyses of effectiveness outcomes except the proportion of patients relapse free at 24 months, where it was 40%, indicating moderate heterogeneity. The results reported in section 4.2.5 are those of the random effects analysis; HRs taken from the manufacturer's response to queries and clarifications are presented for disability progression confirmed at three and six months.

4.2.4 Summary of clinical effectiveness data: individual trials

This summary focuses on annualised relapse rate and confirmed disability progression and briefly reports quality of life and MRI outcomes. These outcomes are prioritised for the following reasons:

ARR was the primary outcome in the CONFIRM study, a secondary outcome in DEFINE, and is recommended by the EMA as a measure for relapse.[14]

Disability progression, measured by an increase of ≥1 point in the EDSS confirmed at three months by two years forms the basis for the economic model. This is a secondary outcome in both the DEFINE and CONFIRM studies. Change in EDSS by 1 point (or 0.5 points if baseline is > 5.5) is recommended by the EMA as a measure of disability progression.

Quality of life is identified by the EMA as a relevant outcome to the assessment of efficacy. The EQ-5D is used in the model for the derivation of QALYs.

MRI outcomes were included in the company submission although not in the NICE final scope. These are regarded as providing potentially relevant indications of disease activity in the absence of relapses, and contribute to diagnostic criteria for RRMS[21] and particular treatment indications.[6, 7]

Safety and tolerability data are summarised in section 4.5.

4.2.4.1 Annualised relapse rate

The ARR for dimethyl fumarate versus placebo in DEFINE was 0.47 (95% CI 0.37 to 0.61), indicating a statistically significant 53% reduction in relapse rate. The primary outcome for DEFINE was the proportion of patients relapsing at two years; this also showed a statistically significant benefit of dimethyl fumarate (HR 0.51, 95% CI 0.40 to 0.66). Steroid treated relapses showed a similar benefit (ARR ratio 0.48, 95% CI 0.36 to 0.63).

The ARR for dimethyl fumarate versus placebo in CONFIRM showed a statistically significant 44% reduction in relapse rate (ARR ratio 0.56, 95% CI 0.42 to 0.74). Steroid treated relapses showed a similar benefit of dimethyl fumarate compared to placebo (ARR 0.56, 95% CI 0.42 to 0.76).

4.2.4.2 Disability progression

The DEFINE study showed a statistically significant benefit for dimethyl fumarate compared to placebo for the secondary outcome of risk of disability progression confirmed at three months (HR 0.62, 95% CI 0.44 to 0.87). A pre-planned sensitivity analysis of disability progression sustained for six months showed a non-statistically significant benefit (HR 0.77, 95% CI 0.52 to 1.14).

The CONFIRM study showed a non-statistically significant benefit of dimethyl fumarate compared to placebo for the secondary outcome of disability progression confirmed at three months (HR 0.79,

95% CI 0.51 to 1.19).[1] The pre-planned sensitivity analysis for disability progression sustained for six months showed a hazard ratio of 0.62 (95% CI 0.37 to 1.03).

The manufacturer suggested (CSR) that the lower estimate of efficacy for progression confirmed at three months was driven in part by the low number (compared to DEFINE) of patients in the placebo group with confirmed progression (and a higher proportion of placebo patients with tentative progression whose data were censored). An ad hoc sensitivity analysis was reported (manufacturer's submission P75), which assumed that patients with tentative progression who withdrew or switched medication did have confirmed progression; although this increased the estimate of efficacy it did not substantially alter the estimate of effect.

The results for the two trials are summarised in Table 7.

Table 7: Confirmed disability progression in the DEFINE and CONFIRM trials

	Progression confirmed at 3 months: HR (95% CI)	Progression confirmed at 6 months: HR (95% CI)
DEFINE	0.62 (0.44 to 0.87)	0.77 (0.52 to 1.14)
CONFIRM	0.79 (0.51 to 1.19)	0.62 (0.37, 1.03)

Results for MSFC composite outcomes and individual components were also reported in the manufacturer's submission, as were progression of cognitive deficit and visual function. The MSFC composite score and two of the three components showed statistically significant benefits of dimethyl fumarate over placebo in the DEFINE trial and, non-statistically significant benefits for these with a trend towards a significant benefit for the composite outcome ($p = 0.058$) in the CONFIRM study. Progression of cognitive deficit and visual function showed no significant differences and little numerical difference between DF and placebo in either trial.

4.2.4.3 Comparisons of glatiramer acetate with placebo and dimethyl fumarate:

The comparison of glatiramer acetate versus placebo showed a statistically significant benefit with active treatment for ARR (ARR ratio 0.71, 95% CI 0.55 to 0.93), which is in line with the observed benefits of glatiramer in previous trials.[5] Steroid treated relapses showed a similar benefit of glatiramer acetate compared to placebo (ARR ratio 0.74, 95% CI 0.56, to 0.99).

For three months disability progression confirmed at two years the comparison between glatiramer acetate and placebo showed a hazard ratio of 0.93 (95% CI 0.63 to 1.37). Six month data showed a hazard ratio of 0.87 (95% CI 0.55 to 1.38).[2]

[1] figure taken from CSR: fig 20 in submission shows 0.62 (0.38, 1.00)
[2] Taken from CSR

Post-hoc analyses of dimethyl fumarate versus glatiramer acetate found that the ARR for dimethyl fumarate versus glatiramer showed a statistically significant benefit for dimethyl fumarate (ARR ratio 0.78, 95% CI 0.61 to 1.00). For disability progression the comparison between dimethyl fumarate and glatiramer acetate gave a hazard ratio of [redacted].[3]

Results from the post-hoc comparison of dimethyl fumarate with glatiramer for other outcomes were also presented; these showed non-significant differences with the same direction of effect.

4.2.4.4 Subgroup analyses for key outcomes

Subgroup analyses for the outcomes of ARR, proportion of patients relapsed at two years and disability progression confirmed at three months at two years were reported for both DEFINE and CONFIRM based on the number of relapses in the previous year (≤1 versus ≥2), McDonald criteria (1 versus 2, 3 or 4), prior MS treatment (this included all treatment , not just DMT), baseline EDSS score (≤2.0 versus >2.0), baseline T2 lesion volume (≤median versus >median), inclusion in MRI cohort, and region of recruitment. These analyses were in the protocol for both trials and were pre-specified.

For the DEFINE study the submission described the subgroups as generally consistent with the overall population. Arguably there is a divergence in the effect sizes in some of the subgroups (Table 8), in particular for the treatment naïve versus treatment experienced patients. Divergence of estimates of effect and confidence intervals which did not overlap or overlapped only marginally were suggestive of a larger benefit in patients who had a baseline EDSS score ≤2 compared to those with baseline EDSS >2. There was some indication of a trend to a similar differential effect in disability progression. There a greater benefit in patients who were treatment naïve, compared to those with prior MS therapy including DMT; there was a consistently larger effect of dimethyl fumarate treatment in the naïve patients, however the confidence intervals of the two groups overlapped.

Due to the inherent limitations of subgroup analyses of subgroups it is not appropriate to draw strong conclusions based on these data.

Table 8: Results of subgroup analyses from DEFINE by baseline EDSS score and prior MS treatment status. Based on Table 24 (P88) in manufacturer's submission

	N	ARR for dimethyl fumarate versus placebo: rate ratio (95% CI)†	Proportion of patients relapsed for dimethyl fumarate versus placebo: HR (95% CI)	Disability progression confirmed at three months for dimethyl fumarate versus placebo: HR (95% CI)
Baseline EDSS score				
EDSS score ≤2	414	0.29 (0.20 to 0.44)*	0.35 (0.24 to 0.51)†	0.52 (0.32 to 0.82)
EDSS score >2	403	0.70 (0.50 to 0.98)*	0.71 (0.51 to 0.99)†	0.73 (0.45 to 1.17)

[3] Figure taken from CSR; RR in the submission. 0.83 (95% CI 0.54, 1.23).

Prior MS treatment				
No prior treatment	368	0.33 (0.21 to 0.52)	0.33 (0.21 to 0.52)	0.38 (0.22 to 0.65)
Prior treatment	450	0.61 (0.45 to 0.84)	0.61 (0.45 to 0.84)	0.83 (0.54 to 1.29)

*95% CI for subgroups are non-overlapping
†95% CI for subgroups overlap only marginally

Differences and trends in treatment effects seen in the DEFINE study were not observed in CONFIRM where the prespecified subgroup analyses conducted by baseline EDSS score and previous treatment status, as well as other variables, generally showed consistency of effect between patient subgroups for the outcomes of ARR, proportion of patients relapsed and confirmed disability progression.

[redacted]

Table 9: Results for subgroups approximating the licensed indications for fingolimod and natalizumab. Taken from Table 21 (P66) in draft EPAR

	ARR for dimethyl fumarate versus placebo: rate ratio (95% CI)†	**Disability progression confirmed at three months for dimethyl fumarate versus placebo: HR (95% CI)‡**
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

†Ratio < 1 favours DF
‡Ratio >1 favours DF

4.2.4.5 Quality of life

The VAS global quality of life measure showed a statistically significant benefit with dimethyl fumarate compared to placebo ($p < 0.01$) in both trials, though this was driven by a deterioration in the placebo group rather than substantial improvement in the dimethyl fumarate group. The EQ-5D VAS showed a statistically significant benefit for dimethyl fumarate in DEFINE ($p < 0.001$) but not in

CONFIRM. This effect was driven by a smaller deterioration in the dimethyl fumarate group than in the placebo group. These results are summarised in Table 10.

The SF-36 summary physical health component score showed a statistically significant benefit in the dimethyl fumarate patients compared to placebo in both trials. However the summary mental health component score did not show a significant benefit in either trial. A number of the health domains which make up these summary scores showed small increases for dimethyl fumarate groups and small decreases in placebo groups across both trials, suggesting that treatment with dimethyl fumarate may impact positively on specific components that affect overall quality of life. However these findings are tentative and further research and analysis would be required before robust conclusions could be made.

Table 10: VAS and EQ-5D results for DEFINE and CONFIRM. Taken from Table 23 (P86) in manufacturer's submission.

QoL measure: difference from baseline	DEFINE			CONFIRM		
	Dimethyl fumarate : mean (SD)	Placebo: mean (SD)	P value	Dimethyl fumarate: mean (SD)	Placebo: mean (SD)	P value
VAS	0.4 (20.0)	-4.0 (22.3)	0.0031	0.3 (22.0)	-3.9 (21.2)	0.0003
EQ-5D	0.00 (0.20)	-0.01 (0.20)	0.0910	0.01 (0.21)	0.00 (0.20)	0.1454
EQ-5D VAS	-0.3 (15.7)	-4.2 (17.8)	0.0008	-1.64 (17.9)	-2.39 (17.86)	0.1783
SF-36 PCS	0.50 (7.1)	-1.4 (7.2)	0.0003	0.49 (7.85)	-0.71 (7.43)	0.0217
SF-36 MCS	0.20 (10.1)	-1.1 (8.9)	0.0651	0.45 (9.80)	-0.07 (10.17)	0.1671

4.2.4.6 MRI data

MRI outcomes were reported for the MRI cohort of patients who comprised approximately 50% of the patients in CONFIRM and 40% in DEFINE. Statistically significant benefits were seen for the numbers of T1, T2 and Gd+ lesions in both trials. Results for these outcomes are summarised in Table 11. These treatment effects represented lower numbers of T1 (hypointense, permanent), T2 (hyperintense, total lesion load) and gadolinium enhancing (active inflammation) lesions on assessment at two years in patients treated with dimethyl fumarate compared to placebo.

Table 11: Key MRI results for DEFINE and CONFIRM. Taken from Table 20 (P78) and summary on P69 of the manufacturer's submission.

MRI measure (lesions):	DEFINE	CONFIRM
	Dimethyl fumarate BID versus placebo: mean ratio (95% CI)	Dimethyl fumarate BID versus placebo: mean ratio (95% CI)
T1	0.28 (0.20 to 0.39)	0.43 (0.30 to 0.61)
T2	0.15 (0.10 to 0.23)*	0.29 (0.21 to 0.41)*
Gd+	0.10 (0.05 to 0.22)	0.26 (0.15 to 0.46)

*some discrepancy with results on p 79 of submission

4.2.5 Summary of clinical effectiveness data: meta-analysis of trials of dimethyl fumarate

The pooled estimate of the ARR of dimethyl fumarate versus placebo from DEFINE and CONFIRM showed a statistically significant benefit [redacted] The estimate for steroid-treated relapses was identical.

The pooled hazard ratio of disability progression sustained for three months also indicated a statistically significant benefit of dimethyl fumarate [redacted]; the pooled estimate for progression sustained for six months was also statistically significant in favour of dimethyl fumarate [redacted]

Pooled estimates were also presented for dimethyl fumarate versus placebo for the following additional efficacy outcomes (see Table 12). These consistently indicated statistically significant benefits of dimethyl fumarate over placebo for the prevention of relapses and sustained increases in disability, assessed by EDSS score. Hazard ratios for time to relapse were not presented.

Table 12: Outcomes related to relapse and disability progression: pooled analyses. Based on Table 26 (P108) and Figures 12-20 (P103-P107) in the manufacturer's submission.

Outcome	Measure	Estimate of effect (95% CI)	I^2 (%)
ARR	Rate ratio	[redacted]	[redacted]
ARR (steroid treated)	Rate ratio	[redacted]	[redacted]
Proportion of patients with relapse at 12 months	RR	[redacted]	[redacted]
Proportion of patients with relapse at 24 months	RR	[redacted]	[redacted]
Proportion of patients relapse free at 12 months	RR	[redacted]	[redacted]
Proportion of patients relapse free at 24 months	RR	[redacted]	[redacted]
Disability progression sustained for 3 months	HR	[redacted]	[redacted]
Disability progression sustained for 6 months	HR	[redacted]	[redacted]
Change in EDSS score at 24 months	WMD	[redacted]	[redacted]

*Taken from manufacturer's response to queries and clarifications 05/07/13, figures 5 and 6

Trials of dimethyl fumarate not included in the submission

The list of excluded studies was supplied by the manufacturer showed appropriate reasons for exclusion, although it included the primary publications of the DEFINE and CONFIRM studies which the ERG presumes to be an error.

The ERG checked the draft EPAR supplied by the manufacturer, and relevant FDA documentation for further studies. The ERG's information specialist also searched MEDLINE and CENTRAL without date restrictions.

Trials not included in the submission which are of particular relevance were study 109MS303 and the Kappos (2008) study of dimethyl fumarate at non-licensed doses only.[22, 23] These were identified from the list of excluded studies for the review of dimethyl fumarate and the draft EPAR supplied by the manufacturer. The ERG accepts the appropriateness of the decision not to include Kappos et al which did not assess the dose for which dimethyl fumarate received a CHMP positive opinion. [23]

However while study 109MS303 does not meet the inclusion critieria, as there is no placebo or active comparator at licensed dose, the ERG felt that it had the potential to provide additional relevant, longer term data, as it was an ongoing study of dimethyl fumarate and was included in the data considered by the EMA (draft EPAR). This study represented a continuation of the DEFINE and CONFIRM trials with all participants re-randomised to dimethyl fumarate arms (either 240mg BID or 240mg TID). The primary objective of the study is to evaluate the long-term safety profile of dimethyl fumarate.

The ERG therefore requested any more recent analyses of data from this trial (the last available data analysis in the draft EPAR had been undertaken in August 2011). The manufacturer responded by referring to a statement in the submission (p156 manufacturers submission), referencing a conference presentation of the most recent safety data from this trial, which stated that the favourable safety profile was maintained in the year subsequent to completion of DEFINE and CONFIRM. The reference cited by the manufacturer in the submission was a report of MRI subgroup data from the CONFIRM study. The ERG identified the correct reference and confirmed that the safety profile of dimethyl fumarate continued to be favourable in an analysis of 1,002 patients with 1,960 patient-years of follow up.[22] Adverse events, adverse events leading to discontinuation and serious adverse events all occurred at rates lower than seen in DEFINE and CONFIRM. Rates in patients switching from placebo or glatiramer acetate to dimethyl fumerate were reported separately and also showed this pattern. Patients switching to dimethyl fumerate reported GI events and flushing as among the most common adverse events. MS relapse and nasopharyngitis also occurred at rates > 10% in all patients. Three deaths occurred none of which were considered related to the study treatment. Fourteen malignancies were also reported.

The manufacturer also stated that an

[redacted]

Efficacy endpoints in 109MS303 include ARR, proportion of patients relapsed, disability progression as measured by EDSS and MRI measures of disease activity at selected sites.

[redacted]

[redacted]

Table 13: Summary of efficacy data from study 109MS303. Taken from Figures 19 and 20 (P76) in draft EPAR.

Group in CONFIRM/DEFINE	Group in 109MS303	Proportion with relapse at 120 weeks: RR (95% CI)	Proportion with EDSS progression confirmed at 24 weeks at 120 weeks: RR (95% CI)
Dimethyl fumarate 240mg BID	[redacted]	[redacted]	[redacted]
Placebo	[redacted]	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]	[redacted]

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

4.3.1 Studies included in the MTC

Trials included in the MTC addressed all of the relevant comparators identified in the NICE scope: interferon beta-1a (Avonex, Rebif 22µg, Rebif 44 µg), interferon beta-1b (Extavia/Betaferon), glatiramer acetate, fingolimod and natalizumab with the addition of teriflunomide for the stated purpose of strengthening the placebo arm.

As may be expected with a network of trials assessing nine interventions in trials conducted over a 20 year time span, there was clinical and methodological heterogeneity between the included studies. Baseline characteristics and key inclusion criteria are shown in Table 14. This table is based on information presented in Table 29 (P115) and Table 30 (P122) of the manufacturer's submission. Only study arms included in the MTC are shown. Although a covariate analysis is presented in the submission exploring the impact of some potential sources of heterogeneity on the MTC results, discussion of heterogeneity was fairly limited in the submission. Similarities and differences between the trials in the MTC are discussed below.

Table 14: Summary of characteristics of trials included in the MTC. Based on Table 29 (P115) and Table 30 (P122) in the manufacturer's submission.

Study	Duration	Treatment arms	N	Key Inclusion criteria (EDSS, baseline relapses, prior treatment history, other relevant criteria)	Mean or median age (years)	Mean or median disease duration (years)	Percentage female	Relapses in previous year (mean or median)	Relapses in previous 2 years (mean or median)	Mean EDSS score
AFFIRM	24 months	Natalizumab 300mg, 4 weekly	627	EDSS score 0 to 5.0 MRI showing lesions consistent with multiple sclerosis with ≥ 1 medically documented relapse within 12 months	35.6	5	72	1.5	NR	2.3
		Placebo	315		36.7	6	67	1.5	NR	2.3
BECOME	24 months	Glatiramer acetate 20mg OD	39	EDSS score 0–5.5 RRMS with ≥ 1 clinical and/or MRI attacks during the 6 previous months OR CIS characteristic of CNS demyelination confirmed on examination; onset within 6 previous months plus evidence of dissemination in time and space	36	1.2	64	NR	NR	2
		Betaferon 250 µg EOD	36		36	0.9	75	NR	NR	2
BEYOND	≥ 24 months (max 3.5 years)	Glatiramer acetate 20mg OD	448	EDSS score 0–5 RRMS with ≥ 1 relapse in the year before entry into the study Treatment-naive	35.2	5.1	68	1.6	NR	2.3
		Betaferon 250 µg EOD	897		35.8	5.3	70	1.6	NR	2.4
BRAVO	24 months	Avonex 30µg weekly	447	RRMS with EDSS 0–5.5 ≥1 relapse in prior 12 months, or 2 relapses in prior 24 months, or 1 relapse in past 12–24 months with 1 Gd+ lesion in the year prior to screening	NR	NR	NR	NR	NR	NR
		Placebo	450		NR	NR	NR	NR	NR	NR
Bornstein 1987	24 months	Glatiramer acetate 20mg OD	25	Aged 20–35 years MS with ≥ 2 exacerbations in the two years before admission Kurtzke Disability Status Scale ≤ 6	30	4.9	56	NR	3.8	2.9
		placebo	25		31	6.1	60	NR	3.9	3.2
Calabrese 2011	24 months	Glatiramer acetate 20mg OD	48	EDSS score ≤ 5.0	38.9	5.5	72.9	NR	NR	2.1
		Avonex 30µg	47		34.8	5.3	68	NR	NR	1.9

Study	Duration	Treatment arms	N	Key Inclusion criteria (EDSS, baseline relapses, prior treatment history, other relevant criteria)	Mean or median age (years)	Mean or median disease duration (years)	Percentage female	Relapses in previous year (mean or median)	Relapses in previous 2 years (mean or median)	Mean EDSS score
		weekly								
		Rebif 44 µg TIW	46		35.9	5.7	69.5	NR	NR	1.9
CONFIRM	24 months	Dimethyl fumarate 240 mg BID	359	EDSS score 0–5.0 ≥ 1 relapse in the 12 months prior to randomisation with a brain MRI demonstrating lesions consistent with MS or evidence of Gd+ lesions of the brain on an MRI performed within the six months prior to randomisation No IFNβ or GA treatment within 3 months of randomisation	37.8	4.9	68	1.3	NR	2.6
		Dimethyl fumarate 240 mg TID	345		37.8	4.6	72	1.4	NR	2.5
		Glatiramer acetate 20mg OD	350		36.7	4.4	71	1.4	NR	2.6
		Placebo	363		36.9	4.8	69	1.4	NR	2.6
Copolymer 1 MS trial	24 months	Glatiramer acetate 20mg OD	125	EDSS score 0–5.0 MS with ≥ 2 relapses in the 2 years prior to entry and onset of the first relapse ≥ 1 year before randomisation	34.6	7.3	70.4	NR	2.9	2.8
		Placebo	126		34.3	6.6	76.2	NR	2.9	2.4
DEFINE	24 months	Dimethyl fumarate 240 mg BID	410	EDSS score 0–5.0 ≥ 1 relapse within the 12 months prior to randomisation, with a prior brain MRI demonstrating lesions consistent with MS, or evidence of Gd+ lesions on an brain MRI performed within the six weeks prior to randomisation No IFNβ or GA treatment within 3 months of randomisation	38.1	5.6	72	1.3	NR	2.4
		Dimethyl fumarate 240 mg TID	416		38.8	5.1	74	1.3	NR	2.4
		placebo	408		38.5	5.8	75	1.3	NR	2.5
Etemadifir 2006	24 months	Avonex 30µg weekly	30	EDSS score 0–5.0 Relapsing MS with ≥ 2 relapses within the 2-year period to treatment initiation	28.1	2.9	80	2	NR	1.9
		Rebif 44 µg TIW	30		27.4	3	76.7	2.4	NR	2.1
		Betaferon 250 µg EOD	30		29.9	3.7	70	2.2	NR	1.9

Study	Duration	Treatment arms	N	Key Inclusion criteria (EDSS, baseline relapses, prior treatment history, other relevant criteria)	Mean or median age (years)	Mean or median disease duration (years)	Percentage female	Relapses in previous year (mean or median)	Relapses in previous 2 years (mean or median)	Mean EDSS score
European and Canadian glatiramer trial	9 months	Glatiramer acetate 20mg OD	119	RRMS for ≥ 1 year EDSS score 0–5 ≥ 1 documented relapse in previous two years and ≥ 1 Gd+ lesion on their screening brain MRI GA naïve	34.1	7.9	77	NR	2.8	2.3
		Placebo	120		34	8.3	72.8	NR	2.5	2.4
EVIDENCE	Median 62 weeks followed by 34 weeks crossover	Avonex 30µg weekly	338	IFN-naïve EDSS scores 0–5.5 ≥ 2 exacerbations in the previous 2 years IFNβ naïve	37.4	6.7	74.6	NR	2.6	2.3
		Rebif 44 µg TIW	339		38.3	6.5	74.9	NR	2.6	2.3
FREEDOMS	24 months	Fingolimod 0.5 µg OD	425	Aged 18–55 years EDSS score 0–5.5 RRMS with ≥ 1 relapses in the previous year and ≥ 2 relapses in the previous two years No IFNβ or GA therapy for ≥ 3 months before randomisation	36.6	8	69.6	1.5	2.1	2.3
		Placebo	418		37.2	8.1	71.3	1.4	2.2	2.5
FREEDOMS II	24 months	Fingolimod 0.5 µg OD	358	EDSS score 0–5.5 RRMS with ≥ 1 relapses in the previous year and ≥ 2 relapses in the previous two years IFNβ or GA therapy stopped 3 or more months before randomization and natalizumab at least 6 months prior to randomization	40.6	41.4	76.8	1.4	2.2	2.4
		Placebo	355		40.1	10.6	81.1	1.5	2.2	2.4
IFNB MS	Median 48 months	Betaferon 250 µg EOD2	124	MS ≥ 1 year EDSS 0–5.5 ≥ 2 acute exacerbations in the previous 2 years, with clinical stability for ≥ 30 days before entry	35.2	4.7	69.4	NR	3.4	3
		Placebo	123		36	3.9	71.5	NR	3.6	2.8

Study	Duration	Treatment arms	N	Key Inclusion criteria (EDSS, baseline relapses, prior treatment history, other relevant criteria)	Mean or median age (years)	Mean or median disease duration (years)	Percentage female	Relapses in previous year (mean or median)	Relapses in previous 2 years (mean or median)	Mean EDSS score
IMPROVE	10 months	Rebif 44 µg TIW	120	Aged 18–60 years RRMS for ≥ 12 months EDSS score 0–5.5 ≥ 1 clinical event and Gd+ lesion within six months prior to randomisation	NR	NR	NR	NR	NR	NR
		Placebo	60		NR	NR	NR	NR	NR	NR
INCOMIN	24 months	Avonex 30µg weekly	92	EDSS score 1–3.5 ≥ 2 clinically documented relapses during the preceding two years No prior treatment with IFNβ	34.9	6.7	62	NR	NR	2.0
		Betaferon 250 µg EOD	96		38.8	5.9	69	NR	NR	2.0
Kappos 2011	12 months	Avonex 30µg weekly	54	EDSS 1–6.0 ≥ 2 relapses within 3 years of screening, ≥ 1 of which occurred in the previous year OR ≥ 6 T2 lesions in the year before screening	NR	NR	NR	NR	NR	NR
		Placebo	54		NR	NR	NR	NR	NR	NR
Knobler 1993	36 months	Betaferon 250 µg EOD	6	RRMS ≥ 1 year, ≤ 15 years EDSS score of 0.0-5.5 ≥ 2 exacerbations in the 2 years prior to entry into the study	35.4	4.2	33.33	NR	NR	2.7
		Placebo	7		34.5	7	71.43	NR	4	3.1
MSRCG	26 months	Avonex 30µg weekly	158	RRMS for ≥ 1 year EDSS score 1.0–3.5 ≥ 2 exacerbations in the previous 3 years and no exacerbations for ≥ 2 months at study entry IFNβ naïve	36.7	6.6	75	NR	2.3	2.4
		Placebo	143		36.9	6.4	72	NR	NR	2.3
O'Connor 2006	9 months	Teriflunomide 7mg OD	61	EDSS score ≤6 ≥2 documented relapses in previous 3 years ; ≥1 clinical relapse in the preceding year	40.1	10.3	75.4	1†	NR	2.5
		Teriflunomide 14mg OD	57		40.1	8.5	79	1†	NR	2
		Placebo	61		39.2	8.6	67.2	1†	NR	2.5

Study	Duration	Treatment arms	N	Key Inclusion criteria (EDSS, baseline relapses, prior treatment history, other relevant criteria)	Mean or median age (years)	Mean or median disease duration (years)	Percentage female	Relapses in previous year (mean or median)	Relapses in previous 2 years (mean or median)	Mean EDSS score
PRISMS	24 months	Rebif 22 µg TIW	189	EDSS scores of 0–5.0 ≥ 2 relapses in the preceding 2 years IFNβ naïve	34.8	5.4	67	NR	3	2.5
		Rebif 44 µg TIW	184		35.6	6.4	66	NR	3	2.5
		Placebo	187		34.6	4.3	75	NR	3	2.4
REGARD	24 months	Glatiramer acetate 20mg OD	378	EDSS score 0–5.5 ≥ 1 attack in the preceding 12 months, IFNβ and GA naive	36.8	NR	72	NR	NR	2.3
		Rebif 44 µg TIW	386		36.7	3.7	69	1[†]	NR	2.4
Saida 2012	6 months, re-randomisation then 6 months	Fingolimod 0.5 µg OD	57	≥1 relapse in previous year or 2 or more relapses in previous 2 years or ≥1 Gd+ lesions within 30 days of study commencement	35	8.2	70.2	1.4	2.2	2.3
		Placebo	57		35	8.2	68.4	1.7	2.8	2.1
TEMSO	24 months	Teriflunomide 7mg OD	366	EDSS score ≤5.5 ≥2 relapses in previous 2 years or 1 relapse in year prior to randomisation	37.4	8.8	69.7	1.4	2.3	2.7
		Teriflunomide 14mg OD	359		37.8	8.7	71	1.3	2.2	2.7
		Placebo	363		38.4	8.6	75.8	1.4	2.2	2.7
TRANSFORMS	12 months	Fingolimod 0.5 µg OD	431	EDSS score 0–5.5 ≥ 1 relapse during the previous year or ≥ 2 two relapses during the previous 2 years	36.7	7.5	65.7	1.5	2.3	2.2
		Avonex 30µg weekly	435		36	7.4	67.8	1.5	2.3	2.2
Wroe 2005	24 months	Betaferon 250 µg EOD	65	RRMS for ≥ 1 year EDSS score 0–5.5 ≥ 2 relapses within the preceding 24 months IFNβ naïve	35	NR	73.9	NR	2.7	2.9
		Placebo	33		38	NR	72.7	NR	2.5	3.1

4.3.1.1 Inclusion criteria

There were some differences in the inclusion criteria for the trials. Notable examples of this include the BECOME trial which included patients with clinically isolated syndrome, whereas other studies did not include this population.[19] Variations in the criteria for EDSS score were minor: all trials required an EDSS score of between either 0 or 1 and either 5.0, 5.5 or 6.0. An exception was the INCOMIN trial which had a maximum EDSS score of 3.5 at trial entry.[24] A minority of trials (e.g. BEYOND)[20] required that patients be treatment naïve with others requiring that patients be naïve to particular therapies (e.g. glatiramer acetate or interferon beta). Criteria for baseline relapses showed some variations, ranging from requirements for ≥2 relapses in the previous three years to ≥1 relapse in the previous six months; 15 of 27 studies had a requirement for ≥1 relapse in the previous year although some required MRI data as an additional measure or allowed MRI evidence as a substitute for a clinical relapse. Some of the more recent studies had additional criteria relating to MRI test results.

4.3.1.2 Baseline characteristics

In addition to the differences in inclusion criterion and outcome measurement there were also differences in the clinical characteristics of the trial populations at baseline. Mean or median disease duration (a mixture of the two measures was reported) ranged from <1 to over 10 years and the submission noted that this variable was inconsistently defined across the trials, with some dating it from first relapse and others from diagnosis. The mean or median relapse rate in the year prior to study entry ranged from 1 to 2.4 which the ERG feel is likely to be clinically meaningful. The mean EDSS score at baseline ranged from 1.9 to 3.2; a score of 2.0 on the EDSS indicates minimal disability in one functional system and a score of 3.0 indicates moderate disability in one functional system or mild disability in 3-4 functional systems, though for both scores patients are fully ambulatory.

There was also some variation between trials in the mean age of participants which ranged from 27.4 to 40.6 years. Disease duration has a well-documented inverse relationship with relapse rate. However, trials with longer mean or median disease duration demonstrated mean or median relapse rates in the one and two years prior to randomisation which were comparable to studies with lower mean/median durations. This is probably a consequence, at least in part, of the inclusion criteria of the trials which required a minimum number of relapses over one or both of these periods. The impact of this lack of correlation, which would normally be present, on the analyses is uncertain.

The full extent of the variability in use of previous DMTs is also unclear. In the trials of treatment naïve patients it was zero, in CONFIRM it was 29% and in DEFINE it was 40% but it was not possible to establish whether these levels were comparable with other included trials. There is also likely to be variation in both the proportion of patients with previous DMT exposure, and the particular treatments to which they had been exposed. Similarly it was not possible to establish the variability in the proportion of patients with highly active or rapidly evolving disease in the trials.

4.3.1.3 Outcome assessment

There appeared to be some variations between the trials in the way in which relapses, and hence ARR were defined, with some trials allowing the presence of a certain number of T2 lesions to be considered as surrogates for relapse.[25, 26] There were also variations in how disability progression was defined. Some studies defined progression based on a ≥1.0 point increase in the EDSS score, with some trials imposing a requirement of ≥1.5 points from a baseline of 0 or, alternatively, 0.5 from a baseline ≥5.5 points. This does not prevent the combining of EDSS data across trials and comparisons within the analysis, but it is a potential source of differences in the outcome measure and its impact on the results of the network analysis is unclear. Whilst most trials used disability progression sustained for 3 months, a minority used the criterion of progression sustained for 6 months. This selection of alternative measures contributed to gaps in the networks where outcome data for a specific comparison were unavailable due to the time-point at which the available trials assessed data. This is partly compensated for by the presentation of networks for comparisons at multiple time-points (e.g. both three and six months sustained disability progression are presented).

4.3.1.4 Duration

Trial duration was also a source of significant heterogeneity, and represented an additional reason why trials included in the MTC did not contribute to individual network analyses; again this is partly compensated for by the presentation of networks at multiple time points (e.g. the proportion of patients with relapses at both 12 months and 24 months). Trial duration ranged from nine to a median of 48 months. A minority of trials had durations shorter than 12 months and it may therefore not have been appropriate to include these in the assessments of clinical efficacy.

4.3.1.5 Statistical aspects

The manufacturer noted (footnote to Table 32, P132) that ARR was reported differently across trials and that standardised methods were used to calculate ARRs for the MTC. Data for many trials were also reported over the two years prior to randomisation only, meaning that relapse rates for the year prior to randomisation required imputation. There are also likely to be additional sources of heterogeneity in the analyses that are not possible to identify from the data presented in the

submission. For example there were differences between the two dimethyl fumarate trials in how patients were censored for the disability progression analysis. It is possible that similar variations extended across the whole data set.

4.3.1.6 Interventions

There is heterogeneity in the indications for which the included interventions are licensed. As the manufacturer correctly noted in their characterisation of current treatment pathways, beta-interferons and glatiramer acetate have EMA licenses for all RRMS patients, although UK prescribing guidelines indicate that treatment should only be started for patients who have had ≥2 relapses in the previous two years.[8] The trial populations are therefore aligned with the licensing criteria. As previously noted, fingolimod and natalizumab are licensed and recommended by NICE for subgroups of RRMS patients who meet the following criteria:

Fingolimod: 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.

Fingolimod is also licensed for patients with RES disease but is not recommended by NICE for use in this group.

Natalizumab: Patients with RES disease defined as 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. The NICE recommendation for use is in accordance with this section of the license.

Natalizumab is also licensed in patients with high disease activity defined as for the fingolimod license.

The trial populations of AFFIRM (natalizumab versus placebo), FREEDOMS (fingolimod versus placebo) and TRANSFORMS (fingolimod versus Avonex) were broader RRMS populations than the groups for which natalizumab and fingolimod are licensed or recommended by NICE. However, the hazard ratios for fingolimod in the highly active disease subgroup and natalizumab in the RES disease subgroup differ from the hazard ratios for the whole population.[6, 7] Estimates of relative treatment

effect for these therapies versus dimethyl fumerate which have been derived from the MTC using whole trial populations will therefore not be applicable to the subgroups for which fingolimod or natalizumab are relevant comparators. The estimates of relative effectiveness which are relevant to the current decision problem are therefore uncertain.

4.3.1.7 Uncertainty

Whilst the manufacturer's submission reported on many aspects of the included trials, data were missing for some characteristics of several trials. There are also likely to be further sources of heterogeneity which were not apparent from the data presented; the impact of variance in ethnicity, McDonald criteria and other variables is uncertain.

4.3.2 Validity of trials in MTC (risk of bias)

The manufacturer's submission included an assessment of some of the key aspects of risk of bias for all the trials included in the MTC. These comprised randomisation, baseline characteristics, blinding of care providers, participants and outcome assessors, imbalances in drop-outs, selective outcome reporting, and use of an appropriate intention to treat analysis. The principal omission from the assessment was adequacy of allocation concealment, which must therefore be regarded as unclear for all trials except DEFINE and CONFIRM.

All trials were reported as either being at low risk of bias or unreported (unclear) risk of bias for each of these characteristics. Seven trials were at low risk of bias for all items in the assessment, of the remaining 20, six were rated as having an unclear risk of bias on only one criterion and 14 had an unclear risk for two or more criteria.

4.3.3 Trials not included in the submission

The MTC included 27 RCTs. The included trials were checked against the 18 RCTs included in the MTC in the appraisal of fingolimod (TA254), which was considered by the ERG in that appraisal to have identified all relevant trials. All except two of those trials (Saida 2005 and Hurwitz 2008) were included. These RCTs were two-arm trials which compared the licensed dose of interferon beta-1b (Betaferon 250µg) with a higher dose.[27, 28] This therefore did not meet the inclusion criteria for the review which required that trials without a placebo/best supportive care comparator should assess ≥ 2 licensed doses of DMTs. (For the purposes of the MTC this criterion appears reasonable).

The ERG also undertook a search of Medline and CENTRAL to identify trials published/performed since TA254 in order to check that all relevant recent trials had been identified/included. One trial which compared glatiramer acetate alone with interferon beta-1a (Avonex) with a combined treatment arm was identified (CombiRx).[29] Outcome data from this trial were published in March 2013 after the

search date for the manufacturer's systematic review; the ERG assume it was not included in the MTC for this reason. It is unclear what the impact of the addition of this trial to the glatiramer and Avonex nodes of the MTC would be; approximately 25% of the trial population of the trial population of n=1008 would contribute to each node.

Other trials which were identified by the ERG as potentially relevant were clearly ineligible according to the inclusion criteria used by the manufacturer for reasons such as agents used in alternative formulations or doses, only tertiary outcomes such as MRI data were reported, or the population comprised patients who did not yet meet MacDonald criteria.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

4.4.1 Statistical methods

The MTC was conducted using PROC GLIMMIX in SAS; this is known to incorporate lower levels of uncertainty around the mean estimates than WINBUGS. This is particularly true when a random treatment effect is modelled. However, a fixed treatment effect was modelled in the manufacturer MTC analyses which will mitigate some of the limitations of using a frequentist approach. A fixed effect model is likely to be appropriate for most of the networks for which analyses were undertaken as there were insufficient trials with which to estimate a between-study variance. The only network that appears as if it may be suitable for a random effects analysis is the network for the annualised relapse rates (Figure 53, P357 in manufacturer's submission). As no random effects analysis was undertaken, the estimate of the confidence intervals of the relative treatment effects may be slightly underestimated (i.e. too narrow).

Although there is some clinical heterogeneity across the trials included in the MTCs (see section 4.3.1), insufficient numbers of trials for most networks prohibits a statistical exploration of the heterogeneity within the MTCs. A covariate analysis was undertaken in an attempt to explore the heterogeneity; however with so few trials levels of uncertainty were high. Two of the covariates (publication year and study duration) were found to be statistically significant; however, the ERG feels that the high level of precision obtained in this analysis is spurious. Given the number of trials included the ERG would anticipate extremely wide confidence intervals, not very precise significant effects. The results of the covariate analyses presented are reported briefly in section 4.4.3.

A number of networks were presented, one for each outcome of interest. The number of trials informing each of the networks varied, depending on outcomes considered and reported in the included trails. As expected, the summary output statistic for each network/outcome varied dependent on the nature of the data being synthesised, and on the whole seemed appropriate. There were a few errors in the code supplied which initially made it unclear whether the binomial outcomes had been modelled as risk ratios or odds ratios; clarification was sought from the manufacturer. Revised code

was supplied and clarification indicated that these outcomes had been modelled as risk ratios in the original submission. Risk ratios are asymmetric and this has been demonstrated to be capable of generating anomalous results in an indirect comparison. However, some confusion remained regarding the effectiveness outcome for disability progression which, like other binomial outcomes, looked as though it had been modelled as a risk ratio, despite being referred to as a hazard ratio. Following ERG queries the manufacturer provided new code and results pertaining to disability progression. The new results were presented as hazard ratios although it is not clear that the code/analysis is appropriate as time does not appear to be included in the model. Given that hazard ratios represent instantaneous risk over the study time period it is unclear to the ERG that these outcomes can be modelled without the element of time. It is the opinion of the ERG that either the code supplied is incorrect or the outcomes modelled are in fact rate ratios. The ERG considers risk ratios or hazard ratios to be preferable to rate ratios for this analysis. Due to this lack of clarity the ERG have opted to use the base case results and model in the manufacturer's submission, which used a risk ratio.

4.4.2 Effectiveness data from the MTC

MTC results were presented for: relapse-related outcomes (ARR, steroid-treated ARR, proportion of patients with relapse at 12 months and at 24 months, and the proportion free from relapse at 12 and 24 months); confirmed disability progression sustained for 3 months and for 6 months at 24 months and change in EDSS score at 24 months. Discontinuation due to any cause and discontinuation due to death, analyses for any adverse event, any serious adverse event and any GI disorder were also presented together with those for a range of individual adverse events (see section 4.5).

The results of analyses for ARR (which is the principal measure of relapse presented in the submission; the primary outcome of CONFIRM and a secondary outcome of DEFINE) and disability progression confirmed at three and six months at two years are presented below (Table 15 and Table 16). Progression confirmed at three months was key secondary outcomes of both CONFIRM and DEFINE and key drivers of the economic model. Progression confirmed at six months was assessed as a sensitivity analysis in the dimethyl fumarate trials but as the main measure of disability progression in a minority of trials of comparator DMTs. Six month sustained progression is regarded by the EMA and the ERG's clinical advisor as a more robust measure than three month sustained progression.[29]

Although 27 trials were included in the MTC, not all of these contributed data on every outcome. The network for ARR is the most complete, incorporating data from all except one trial (Wroe 2005).[30] Nine trials contributed data for disability progression sustained for three months at 24 months and, although all comparators except Avonex were represented in this network, the majority of comparisons were populated by single trials. Inclusion of studies in the network for confirmed

disability progression was limited by the fact that only confirmed progression by 24 months was assessed. Hence studies with shorter durations such as the TRANSFORMS and EVIDENCE trials which assessed comparisons with Avonex were not included. In general, the networks for outcomes not related to relapse occurrence were relatively sparsely populated, with most comparisons represented by single trials and some comparisons/comparators not assessed at all. This was particularly the case for individual adverse events.

4.4.2.1 Relapse rate

The ARR comparisons demonstrated that dimethyl fumarate was statistically significantly more effective than placebo, all interferon regimes and glatiramer acetate and teriflunomide. Natalizumab was statistically significantly more effective than dimethyl fumarate, while fingolimod showed a trend towards increased effectiveness compared to dimethyl fumarate. Results for steroid treated relapses showed statistically significant benefits against Avonex and Rebif 22µg as well as placebo. Non-significant benefits were observed in comparisons with other traditional DMT regimes (Rebiff 44µg, Betaferon, glatiramer acetate) and results for natalizumab and fingolimod showed the same pattern as those for ARR. The network for steroid-treated relapses was more sparsely populated (12 trials) than that for the main ARR analysis and did not include teriflunomide.

Table 15: Results of MTC analysis for ARR taken from Figure 21 (P139) in manufacturer's submission.

	ARR rate ratio (95% CI)
Placebo	[redacted]
Glatiramer acetate	[redacted]
Avonex	[redacted]
Betaferon	[redacted]
Rebif 22µg	[redacted]
Rebif 44µg	[redacted]
Fingolimod	[redacted]
Natalizumab	[redacted]
Teriflunomide 7 mg	[redacted]
Teriflunomide 14 mg	[redacted]

4.4.2.2 Proportions of patients with disability progression confirmed at three and six months

Confirmed disability progression sustained for three months at 24 months showed a statistically significant benefit for dimethyl fumarate only when compared to placebo (Table 16). There were no other statistically significant differences between dimethyl fumarate and the other comparators. The direction of effect favoured dimethyl fumarate for all comparisons except that with natalizumab. Analysis of confirmed progression sustained for six months did not show a statistically significant difference between dimethyl fumarate and any comparator, including placebo (also Avonex,

Betaferon, Rebiff 44µg, glatiramer acetate and fingolimod). This indicates a further need for caution in considering the results of the direct comparison between dimethyl fumarate and placebo for three months sustained progression. The analysis of mean change in EDSS score from baseline at 24 months also indicated no statistically significant difference between dimethyl fumarate and any other comparator including placebo (all comparators except terifluonmide were represented). This network was less sparse than that for progression confirmed at six months. As with the meta-analysis of dimethyl fumarate compared to placebo, the submission presented relative risks rather than hazard ratios for this outcome. In response to queries raised by the ERG the manufacturer supplied a revised network. This appeared to give rate ratios rather than hazard ratios. This did not show large differences in effect sizes compared to the analyses presented in the original submission, and did not alter the results consistently in favour of either comparators or dimethyl fumarate.

Table 16: Results of MTC analysis for EDSS 3 and 6 months confirmed disability progression at 24 months. Based on Figure 28 (P145) and Figure 29 (P146) in manufacturer's submission.

	EDSS progression confirmed at three months: relative risk (95% CI)	EDSS progression confirmed at six months: relative risk (95% CI)
Placebo	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]
Avonex	[redacted]	[redacted]
Betaferon	[redacted]	[redacted]
Rebif 22µg	[redacted]	[redacted]
Rebif 44µg	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]
Natalizumab	[redacted]	[redacted]
Teriflunomide 7 mg	[redacted]	[redacted]
Teriflunomide 14 mg	[redacted]	[redacted]

4.4.3 Covariate analyses

The submission included both univariate and multivariate analyses of a number of covariates in the MTC for the outcomes of ARR, EDSS progression, at three and six months, proportion of patients relapse free at 24 months and discontinuation for any reason. The ERG's concerns about the power of covariate analyses in networks containing limited numbers of trials are discussed in section 4.4.1.

The covariates assessed included the majority of variables identified by the ERG as sources of clinical heterogeneity between the trials included in the MTC. Variables assessed were: study duration, mean age, mean disease duration, percentage of female participants, relapses in year prior to randomisation, mean EDSS score at baseline and publication year.

In the both univariate and multivariate analyses for ARR the only significant covariates were publication year and relapse in the one year prior to randomisation. Estimates of effectiveness

adjusted for both significant and nonsignificant covariates were provided; these found that the estimates of effect for individual treatment comparisons were only minimally impacted by adjustment.

Univariate analyses were undertaken for sustained disability progression confirmed at both three and six months which identified no significant covariates; the multivariate models did not converge due to insufficient contributing studies. As with ARR, only minimal differences in estimates of effect for individual treatment comparisons were seen.

4.5 Tolerability and safety

There were three sources of information on the tolerability and safety of dimethyl fumarate in the manufacturer's submission. These were

(i) Summary of individual adverse event incidence from the DEFINE and CONFIRM trials (manufacturer's submission, Table 37 (P155),

(ii) Pooled risk ratios for adverse events from the DEFINE and CONFIRM trials (manufacturer's submission, Table 27 (P109)),

(iii) MTC for adverse events from the DEFINE and CONFIRM studies; risk ratios for adverse events for dimethyl fumarate versus placebo and active comparators (manufacturer's submission, Table 35 (P148)).

Tolerability and safety outcomes for the comparison of dimethyl fumarate with placebo were reported individually for DEFINE and CONFIRM but comparisons between groups were reported only in a pooled analysis of the DEFINE and CONFIRM trials (see below). Although there were differences between the two trials the ERG regards this as reasonable and appropriate. Statistical heterogeneity between the trials was low for a majority of outcomes. $I^2 = 0\%$ in a majority of cases. I^2 exceeded 40% for the following outcomes: ALT increased (87%); diarrhoea (41%); pain in extremity (56%); discontinuation due to adverse effects (44%).

The submission did not include a comparison of the pooled dimethyl fumarate arms with glatiramer acetate; data were reported for the placebo versus glatiramer acetate comparison from the CONFIRM trial. However the MTC included analyses of overall tolerability and safety outcomes as well as some specific adverse events.

The manufacturer defined common adverse events as those which occurred in ≥5% of patients in the DEFINE and CONFIRM trials. Adverse event results were reported in the submission regardless of statistically significant differences between treatment groups, which was appropriate. Statistically significant results were highlighted in tables. This was reasonable although there was limited consideration of adverse events which showed trends of increased incidence in dimethyl fumarate

groups. For the purposes of summary, the ERG briefly summarise mainly statistically significant differences in adverse effects but also note some trends.

Not all outcomes which occurred in ≥5% of patients in DEFINE and CONFIRM were included in the pooled analysis or calculation of relative risk in the case of comparison with glatiramer acetate. Visual inspection of the incidence rates for these outcomes by the ERG indicated that, with the exception of hot flush, they did not include events for which differences were likely to be statistically significantly higher in the dimethyl fumarate groups. The submission acknowledged the importance of hot flush as an adverse event. Notable among those events which were not included in the pooled analyses was MS relapse which was included as an adverse event but was appropriately assessed in the efficacy analyses. Other events were captured by pooled analyses of groups of events such as lower respiratory tract infections. Renal monitoring is required for dimethyl fumarate, but individual renal outcomes (haematuria, microalbuminuria, proteinuria) did not show substantive differences between the groups in DEFINE and CONFIRM (haematuria was reported in DEFINE only). Other events such as hypoaesthesia and parasthesia also occurred at equivalent rates across all groups. Pooled estimates of effect were reported for two tolerability outcomes. These indicated that there were higher levels of discontinuation due to adverse events in patients in dimethyl fumarate groups than in placebo groups ([redacted]), but that discontinuation due to any cause was lower in dimethyl fumarate groups ([redacted]). Discontinuation due to death was reported only for the CONFIRM trial, but favoured dimethyl fumarate ([redacted]).

Analyses showed that dimethyl fumarate and glatiramer acetate did not differ significantly for discontinuation due to adverse events ([redacted]), discontinuation due to any cause ([redacted]) or discontinuation due to death which favoured dimethyl fumarate ([redacted]).

Pooled analyses were also presented for occurrence of any adverse event, any serious adverse event and any gastrointestinal (GI) disorder, as well as for the individual adverse events assessed in the DEFINE and CONFIRM studies. Any GI disorder was statistically significant higher in the dimethyl fumarate groups (Table 17). Both all adverse events and all serious adverse events occurred at rates comparable to placebo.

There was no increased risk of infectious disease in dimethyl fumarate groups compared to placebo in either trial. The submission stated that decreases in white blood cell and lymphocyte counts were observed but that mean values of these remained within normal limits; further details reported in the CSRs of CONFIRM AND DEFINE confirmed this. The ERG noted a trend in the pooled analysis towards increased incidence of leukopenia in dimethyl fumarate groups versus placebo (RR 4.88, 95% CI 0.84 to 28.32).

The ERG is aware of four cases of progressive multifocal leukencephalopathy (PML) in patients treated with fumaric esters.[31-33]

[redacted] The ERG's clinical advisor stated that, if PML were established as an adverse effect of dimethyl fumarate, guidance for discontinuation based on lymphocyte counts would be required and that monitoring of these parameters would therefore be required in clinical practice.

Individual adverse events which were statistically significantly more common in dimethyl fumarate than in placebo groups are shown in Table 17 and events which were significantly more common in dimethyl fumarate compared to glatirmaer acetate in Table 18. Injection site erythema occurred in 9% of patients in the glatiramer acetate arm, and injection site pain in 8%; these events were not applicable to patients in dimethyl fumarate or placebo groups.

Table 17: Adverse events which occurred statistically significantly more often in dimethyl fumarate groups versus placebo (pooled analysis). Based on Table 27 (P109) in manufacturer's submission.

Outcome	DEFINE		CONFIRM		Pooled RR (95% CI) (random effects)	I^2 (%)
	Dimethyl fumarate (N =410)	Placebo (N = 408)	Dimethyl fumarate (N = 359)	Placebo (N = 363)		
Abdominal pain	46 (11)	22 (5)	27 (8)	15 (4)	[redacted]	[redacted]
Upper abdominal pain	40 (10)	28 (7)	36 (10)	17 (5)	[redacted]	[redacted]
Nausea	53 (13)	38 (9)	27 (8)	15 (4)	[redacted]	[redacted]
Any GI disorder	NR	NR	NR	NR	[redacted]	[redacted]
Flushing	154 (38)	20 (5)	110 (31)	13 (4)	[redacted]	[redacted]
Pruritis	42 (10)	19 (5)	20 (6)	11 (3)	[redacted]	[redacted]
Rash	34 (8)	13 (3)	24 (7)	13 94)	[redacted]	[redacted]

Table 18: Adverse events for which there were statistically significant differences between dimethyl fumarate and glatiramer acetate (CONFIRM trial).* Based on Table 27 (P109) in manufacturer's submission.

Outcome	Dimethyl fumarate (N = 359) N (%)	Glatiramer acetate (N = 351) N (%)	RR (95% CI)
Any adverse event	338 (94)	304 (87)	[redacted]
Any GI disorder	NR	NR	[redacted]
Abdominal pain	27 (8)	4 (1)	[redacted]
Upper abdominal pain	36 (10)	4 (1)	[redacted]
Nausea	40 (11)	15 (4)	[redacted]
Diarrhoea	45 (13)	14 (4)	[redacted]
Flushing	110 (31)	6 (2)	[redacted]
Pruritus	20 (6)	7 (2)	[redacted]
Rash	24 (7)	8 (2)	[redacted]

*all occurred more often in dimethyl fumarate group

As can be seen from Table 17 and Table 18, the adverse events which were significantly more common in dimethyl fumarate groups fell into three categories: GI events, skin disorders and flushing.

Analysis provided in the submission indicated that most GI and flushing events occurred in the first three months of treatment with dimethyl fumarate with many fewer in the subsequent three months, and they continued to diminish over time thereafter. This pattern was seen in both the DEFINE and CONFIRM studies. No comparable data were provided for skin disorders.

The MTC assessed a range of individual adverse events as well as discontinuations and summary measures of all adverse events, all serious adverse events and all GI events. The results of the MTC for summary event categories are shown in Table 19 and Table 20; these represent the networks to which most trials contributed and which provide an overview of the toxicity of the different therapies. GI events are discussed because this category shows an excess of occurrences in dimethyl fumarate which was significant in the direct comparisons. Those individual events which occurred significantly more often in dimethyl fumarate groups are also briefly discussed.

The MTC found no statistically significant differences in discontinuation for any cause between dimethyl fumarate and any of the comparators (all comparators were represented) (Table 19). There was no consistent direction of effect in the analyses. The comparison with glatiramer acetate mirrored the direct comparison in the CONFIRM trial in favouring glatiramer. Discontinuation due to death showed a similar lack of statistically significant differences between dimethyl fumarate and comparators but did not include estimates for the comparison with teriflunomide. Covariate analysis for the factors identified in (section 4.4.3) was undertaken. Publication year and study duration were significant covariates in the univariate analysis (both positively correlated with outcome) but no variables were significant in the multivariate analysis. Analyses of effect estimates adjusted for all variables based on both models were provided and indicated that impact of adjustment was minor.

Table 19: Results of MTC analyses for discontinuations of treatment. Based on Figure 30 (P151) and Figure 31 (P152) in the manufacturer's submission.

	Discontinuation for any cause: RR (95% CI)	Discontinuation due to death: RR (95%CI)
Placebo	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]
Avonex	[redacted]	[redacted]
Betaferon	[redacted]	[redacted]
Rebif 22µg	[redacted]	[redacted]
Rebif 44µg	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]
Natalizumab	[redacted]	[redacted]
Teriflunomide 7 mg	[redacted]	[redacted]
Teriflunomide 14 mg	[redacted]	[redacted]

The MTC included analyses for any adverse event, any serious adverse event and any GI disorder, and for the individual adverse events assessed in the DEFINE and CONFIRM studies, but not for discontinuations due to adverse events. Results of the MTC for dimethyl fumarate versus active comparators and placebo for any adverse event, any serious adverse event and any GI event are shown in Table 20. No comparison with Betaferon was possible and GI events were only available for the comparisons assessed in the dimethyl fumarate trials (placebo and glatiramer acetate). Few statistically significant differences were observed although there were more adverse events of any kind compared to glatiramer acetate, and significantly fewer than with Rebif 44µg.

Table 20 Results of MTC analyses for categories of adverse events. Based on Table 35 (P148) in the manufacturer's submission.

Treatment	Any adverse event: RR (95% CI)	Any serious adverse event: RR (95% CI)	Any GI disorder: RR (95% CI)
Placebo	[redacted]	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]	[redacted]
Avonex	[redacted]	[redacted]	[redacted]
Rebif 22µg	[redacted]	[redacted]	[redacted]
Rebif 44µg	[redacted]	[redacted]	[redacted]
Betaferon	[redacted]	[redacted]	[redacted]
Natalizumab	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]
Teriflunomide 7mg	[redacted]	[redacted]	[redacted]
Teriflunomide 14 mg	[redacted]	[redacted]	[redacted]

Individual adverse events showed a similar pattern to those seen in the pooled analyses from the DEFINE and CONFIRM studies above; flushing, GI disorders, rash and pruritus again occurred more often compared to both placebo and glatiramer acetate although not all differences were statistically significant; nausea was also seen at higher levels compared to interferons. Increased levels of alanine transferase were significantly less frequent compared to Rebif 22µg, Rebif 44µg, Betaferon and fingolimod.

4.6 Conclusions of the clinical effectiveness section

The submitted evidence closely accords with the decision problem defined in the scope. The most important difference between the scope and the submitted analyses is that fingolimod and natalizumab are relevant comparators only for the subgroups for which they are licensed and recommended by NICE, but they are included in the MTC as comparators for all RRMS patients. The ERG recognises the rationale for this, because these treatments were evaluated in RCTs in general RRMS populations, but notes that licencing and guidance decisions were based on estimates of effect from the subgroups and not the whole trial populations. An additional comparator, teriflunomide, was also included in the MTC; the manufacturer's rationale for it having been licensed by the FDA. Another deviation from

the scope was the fact that severity of relapse was not included as an outcome because this was not assessed in the trials of dimethyl fumarate.

Included trials directly assessing dimethyl fumarate were good quality phase III placebo-controlled RCTs at low risk of bias. A pooled analysis of these studies used appropriate methods.

Both the DEFINE and CONFIRM studies showed reduced annualised relapse rates in patients treated with dimethyl fumarate relative to placebo; CONFIRM also showed a benefit for dimethyl fumarate compared to glatiramer acetate although this was a post-hoc analysis. These benefits were reflected in the pooled analyses and in the mixed treatment comparison.

The results for disability progression were somewhat more equivocal. For three month confirmed disability there was a statistically significant benefit in the DEFINE trial for dimethyl fumarate versus placebo. However, in the CONFIRM trial, although the hazard ratio was in the direction of benefit from dimethyl fumarate, the confidence intervals included there being no benefit. Pooled analysis showed a statistically significant benefit. Three month confirmed disability progression has limitations as an outcome measure although it is commonly used in clinical trials. Six month confirmed progression, which is a more reliable measure of permanent disability progression was also presented; although benefits were seen in individual trials the confidence intervals in both cases included there being no benefit of dimethyl fumarate. Pooled analysis did demonstrate a statistically significant benefit. Quality of life and MRI measures also showed evidence of benefit over placebo.

The MTC used to assess effectiveness compared to other comparators identified in the NICE scope identified all the relevant trials. These trials all appeared to be at low or unclear risk of bias although allocation concealment was not assessed. As with the pooled analysis of DEFINE and CONFIRM, relative risks were presented for disability progression, rather than hazard ratios. Some networks were sparsely populated due to the combination of outcomes selected for analysis and the available data from the included trials. There was a moderate level of clinical and methodological heterogeneity between trials in the MTC. This included variations in baseline characteristics such as mean EDSS score and in inclusion criteria including number of relapses in the period prior to randomisation. There was also heterogeneity between the interventions assessed: beta-interferons and glatiramer acetate are licensed/recommended for broad RRMS populations which meet a requirement for two relapses in two years but natalizumab and fingolimod are licensed only for the RES and highly active disease subgroups and recommended by NICE only for these groups respectively. The impacts of some of these differences on the network were assessed for key effectiveness outcomes using multivariate and univariate analyses of covariates, although these may not be reliable due to the small numbers of trials in the disability progression networks. The impact of other sources of heterogeneity remains unclear.

Safety and tolerability data were presented for the dimethyl fumarate trials, pooled analyses and MTC. These analyses appeared complete with respect to relevant adverse events and indicated few serious adverse events associated with dimethyl fumarate. Higher incidences of flushing and GI events were reported but appeared largely confined to the first months of treatment; it was unclear whether this was also the case for skin disorders. Incidences of PML associated with fumaric esters have been documented;

[redacted]

The totality of the clinical evidence submitted by the manufacturer appears to provide a complete and unbiased representation of the available evidence for the efficacy and safety of dimethyl fumarate compared to placebo and other available DMT.

However, the submitted evidence is based on trials with a duration of two years. Although this is comparable with other RRMS trials it still represents a short period in the context of the lifelong duration of MS and the substantial duration of RRMS (with 50% of patients still in RRMS ten years after diagnosis). Therefore there is considerable uncertainty as to the long term efficacy of dimethyl fumarate.

The relative effectiveness of dimethyl fumarate compared to fingolimod and natalizumab has not been established in the populations for which these comparator treatments are indicated. It should also be noted that all the available evidence relates primarily to RRMS patients who would meet criteria for currently available DMT. These patients have a slightly higher relapse rate than the whole UK clinical population of RRMS patients.

5 Cost Effectiveness

This section focuses on the economic evidence submitted by the manufacturer in the initial report and the additional information received in response to the ERGs requests for clarification. The submission is subject to critical review on the basis of this evidence and through the direct examination of the electronic version of the economic model. The appraisal will be presented in the form of a narrative, highlighting key assumptions and possible limitations. A checklist will be used to help guide this narrative and provide an indication of the evaluation's quality. Where possible, the issues highlighted are further explored in additional analyses undertaken by both the manufacturer, at the clarification stage, and the ERG. The ERG's exploratory analyses are presented in Section 6.

The manufacturer's initial economic submission to NICE included:

1. A systematic review of the cost-effectiveness evidence for dimethyl fumarate in adult patients with RRMS (P164 to 170);
2. A description of the de-novo economic evaluation conducted by the manufacturer; including details of the intervention; comparators and patient population; the modelling methodology; the resource components and unit costs; data input sources and assumptions; the base case results; and sensitivity analysis (P171 to 274);
3. A *de-novo* economic evaluation including an electronic version of the model in Microsoft Excel.

Following the points of clarification raised by the ERG, the manufacturer provided a response to the points of clarification. The key points for clarification included:

- Request for justification regarding the use of relative risks rather than hazard ratios to inform disease progression estimates used in the model;
- Clarification around the modelling undertaken to obtain EDSS health state costs;
- Justification for the choice of variance parameters used to inform distribution in the probabilistic sensitivity analysis.

A brief outline of the economic elements of the submission is presented first, followed by a more detailed summary and critique. A summary of the manufacturer's approach and signposts to the relevant sections in the manufacturer's submission are reported in Table 21.

5.1 Overview

The manufacturer conducted a review of the cost-effectiveness literature, but this did not appear to inform the manufacturer's choice of model. No economic evaluation including dimethyl fumarate was identified. The choice of model was based on previously validated models produced for NICE and used in previous NICE technology appraisals for drugs for MS.[5, 7]

In the analysis presented by the manufacturer, dimethyl fumarate was compared to Rebif 22µg, Rebif 44µg, Avonex, Betaferon, glatiramer acetate, natalizumab and fingolimod for the general RRMS population. The population evaluated in the model reflected the population in the dimethyl fumarate trials as discussed in section 4.2.2. Although no distinction was made by the manufacturer, the comparators evaluated included both drugs recommended by NICE and licensed for first-line treatment (Rebif 22µg, Rebif 44µg, Avonex, Betaferon, glatiramer acetate) and drugs recommended and licensed for patients with rapidly evolving severe disease or patients with highly active disease (natalizumab, fingolimod). The ERG discusses this issue in sections 5.2.2 and 5.2.3.

A Markov model was presented which characterised the natural history of the disease through patient progression from RRMS to SPMS. Whether a patient has RRMS or SPMS there is a possibility of disability progression, which is characterised by 10 EDSS states, 0 to 9. In addition to disability progression (i.e. get worse) RRMS patients may also regress to lower EDSS states (i.e. improve). SPMS patients cannot regress. The progression to SPMS from RRMS is independent of treatment and the likelihood of progressing varies according to the EDSS state. Patients may die at any time. The mortality rate was assumed equal for both RRMS and SPMS patients.

The drugs in the model affect the health of patients and cost to the health system through reduction in the annual relapse rate, the reduction in the annual risk of disability progression for a patient with RRMS, and through the occurrence of adverse events. Patients with SPMS do not receive treatment. Effectiveness data were obtained from a mixed treatment comparison of trials with a general RRMS population (see section 4.1.2). The time horizon of the model was 30 years. In the manufacturer's base case, the treatment effect on disability progression was assumed to wane after 2 years to 75% of the original effect for the third, fourth and fifth years of treatment, followed by 50% for every remaining year on treatment. Relapse effects and adverse events were assumed constant while on treatment.

Patients may discontinue the drug due to adverse events, by moving to an EDSS state of 7 or higher, or through progression to SPMS. Discontinuation of treatment results in the patient receiving no treatment for the remainder of the model time horizon.

Model outcomes were measured in terms of quality-adjusted life-years (QALYs) based on comparative effectiveness data and health-related quality-of-life (EQ-5D). Utility data were obtained from the pooled data of dimethyl fumarate trials, supplemented by the UK MS survey. A brief description of the MS survey is given in section 5.2. Resource use was also derived from the UK MS survey and unit costs from relevant national source were then applied.

Base case results were presented as pair-wise incremental cost-effectiveness results for dimethyl fumarate versus each of the comparators. In addition, a full incremental analysis was presented where drugs were compared to the next most expensive. The only treatment that produced more QALYs than dimethyl fumarate was natalizumab.

When the list price was used for all drugs, pairwise comparisons showed that dimethyl fumarate was not cost-effective compared to the other comparators with the ICER for dimethyl fumarate ranging from £106,127 to £173,745. However, fingolimod was dominated by dimethyl fumarate, i.e. was more costly and less effective. An incremental analysis using the list price for all drugs showed that glatiramer had an ICER of £15,026 when compared to Rebif 22µg; dimethyl fumarate an ICER of £159,295 compared to glatiramer acetate; and natalizumab an ICER of £173,745 compared to dimethyl fumarate.

When the manufacturer's proposed PAS price was used for dimethyl fumarate while using the list prices for all other drugs, dimethyl fumarate now dominated, i.e. was less costly and more effective, Rebif 44µg, fingolimod, Betaferon and Avonex. The ICER of dimethyl fumarate was £18,581 compared to Rebif 22µg and £19,057 compared to glatiramer acetate. The ICER of natalizumab compared to dimethyl fumarate increased to £534,047. An incremental analysis using the proposed PAS price showed that glatiramer acetate had an ICER of £15,026 when compared to Rebif 22µg; dimethyl fumarate an ICER of £19,057 compared to glatiramer acetate; and natalizumab an ICER of £534,047 compared to dimethyl fumarate.

Deterministic sensitivity analyses generally showed that results were robust to those parameters tested. Probabilistic results using the PAS price suggested that the probability of dimethyl fumarate being cost-effective was roughly 50% at a threshold of £30,000.

Table 21: Summary of the manufacturer's economic evaluation (and signposts to manufacturer's submission)

	Approach	Source/Justification	Location in manufacturer's submission
Population	The population was the general RRMS population, which was the population in the dimethyl fumarate trials.	The general RRMS population is based on the licensed indication for dimethyl fumarate. Although natalizumab and fingolimod are both recommend by NICE for patients with rapidly evolving severe disease or patients with highly active disease, the populations of the trials were patients with general RRMS.	Section 3.5; P50 Section 7.2.2; P171 Section 7.2.7; P173
Comparators	Dimethyl fumarate was compared to treatments recommended for the general population of RRMS patients (Rebif 22µg, Betaferon, Rebif 44µg, Avonex, and glatiramer acetate), and to treatments recommended for rapidly evolving (natalizumab) and highly active (fingolimod) disease.	This was based on the product indications and the relevant NICE guidelines. The selection was also consistent with the scope. Hence, there was no best supportive care comparator.	Section 2.7; P26 Section 5; p.31
Model, states and events	A cost-effectiveness (cost-utility) analysis was undertaken using a Markov model. The Markov model contains 21 states: 10 EDSS states, each for RRMS and SPMS, and one for death. From an RRMS EDSS state there is an annual probability of making a transition to another RRMS EDSS state and to SPMS. From an SPMS EDSS state there is a probability of making a transition to higher SPMS EDSS state. Within each EDSS state, there is an annual risk of relapse.	The model structure was based on a validated model developed for NICE, and which has been used three previous health technology submissions.[5-7] The natural history of the disease is characterised by patients progressing from RRMS to SPMS. Within each of these there is disability progression, which is commonly measured by EDSS score.	Section 7.2.2/3; P171/2
Natural History	Natural history was characterised through annual transition between the following health states: Between EDSS states in the RRMS condition. Between RRMS EDSS states to SPMS EDSS states.	The transition probabilities from EDSS states 0-7 in the RRMS condition to other states were derived from the placebo arms of the DF trials. The transition probabilities from EDSS states 8/9 in the RRMS condition to other states were extrapolated using London Ontario data. All of these data were derived using the Multi State Markov statistical technique.	Section 7.3.2; P184 Section 7.3.2; P185 Section 7.3.1; P176/7

	Between EDSS states in the SPMS condition. The probabilities of relapse and remission within each EDSS state were included. The annual probability of death.	The transition probabilities between RRMS EDSS states to SPMS EDSS states were based on the London Ontario dataset. Transitions between SPMS EDSS states were based on the London Ontario dataset. The annualised relapse rate within each EDSS state was based on the 12 months prior to enrolment in the DF trials for EDSS states 0-5. In the base case, the rates within states 6-9 were extrapolated from the lower states using data from the MS survey. London Ontario or MS survey data used for higher EDSS states due to a low number of patients in the trials for those states. The probability of death was based on a publication by Pokorski as it provided the most conservative mortality multiplier estimates.[34]	Section 7.3.1; P179
Treatment effectiveness	There were two treatment effects in the model: the annual relapse rate; and the relative risk of disability progression as measured by the EDSS scale.	Both treatment effects were estimated using all the direct and indirect trial evidence for all the comparators in mixed treatment comparisons.	Section 6.7; P113-146 Section 7.3.1; P179/80
HRQoL	Utilities were derived for the 10 EDSS states, for both the RRMS and SPMS conditions and with and without relapse. Disutilities were derived for the adverse event states. Disutilities experienced by caregivers for each EDSS state were derived.	In the base case, the utilities for the 10 EDSS states in the RRMS condition with no relapse were derived from the pooled arms of the DF trials. The utility differences between the SPMS and RRMS conditions and between the relapse and no relapse states were estimated from the MS survey. The disutilities associated with adverse events were derived from published sources and expert opinion. The disutilities experienced by caregivers were also estimated from the MS survey data.	Section 7.4.9, P200/201 Section 7.4.8, P199/200
Adverse events	24different adverse events were included. An adverse event was included in the model if either there was an incidence rate of at least 5% in any of the treatment arms; it is a common dimethyl fumarate adverse event on label and extracted in the systematic review;	The annual incidence of adverse events was calculated from the systematic review, using a weighted average across studies. The proportion of each event that was serious was also as calculated using the number of serious events reported in the systematic review.	Section 6.9.2; P68-71 Section 6.9.3; P72 Section 7.3.1; P89-90

	or it is an adverse event occurring at an incidence rate of at least 3% higher in the total dimethyl fumarate group than in the placebo group.		Section 7.4.9; P114-115 Section 7.5.7; P132-133
Resource use and costs	The NHS resource use and costs associated with dimethyl fumarate and its comparators were estimated and included treatment costs, administration costs and costs of monitoring. Treatment costs were estimated by the multiplying doses per year by the unit costs for dimethyl and the comparators. The administration and monitoring costs associated with dimethyl fumarate and its comparators were then included. The health state costs incorporated the EDSS state costs and the average cost of relapse. These were estimated using a seemingly unrelated regression. The adverse event unit costs were applied to the adverse event incidence rates outlined above.	Drug acquisition costs were calculated using the British National Formulary and the doses per year indicated in the drugs' marketing authorisations.[35] Treatment acquisition and monitoring costs were derived from Department of Health reference costs and the National Tariff and resource was derived from product indications.[36, 37] Health-state resource use was estimated from the UK MS Survey and costs were derived from the Department of Health reference costs and the National Tariff. Incidence of adverse events are outlined about and the associated costs were derived from Personal Social Services Research Unit (PSSRU), NHS Reference costs and manufacturers assumptions. [36, 38]	Section 7.5.1; P202 Section 7.5.2-4; P203 Section 7.5.5; P203-206 Section 7.5.6; P206-212 Section 7.5.7; P212-214
Discount rates	Costs and benefits were discounted at 3.5% per annum.	In accordance with the NICE reference case.	Section 7.2.6; P82
Sensitivity analysis	One-way and/or two-way sensitivity analysis were conducted around the treatment waning effect and the annual discontinuation risk Two-way sensitivity analysis was conducted on the relapse rates, disability progression risk ratio and the drop out rates. Several scenario analyses were also undertaken: no treatment waning effect; 0 to 6% discount rates; 1 to 50 year time horizons; mortality rate of MS population equal to the general population; alternative annualised relapse rates; alternative disability progression; London Ontario transition matrix for RRMS-RRMS transitions. Probabilistic sensitivity analysis, using 1,000		Section 7.6; P214-218 Section 7.7.7; P238-267

	iterations, was undertaken for both scenarios where the list price and the manufacturer's proposed PAS price were used for dimethyl fumarate. The outputs of the PSA were mean cost and QALY estimates for each treatment along with 95% confidence intervals; scatterplots; and cost-effectiveness acceptability curves.		

5.2 ERG comment on manufacturer's review of cost-effectiveness evidence

The manufacturer's submission described the search strategies used to identify relevant studies relating to the cost effectiveness of dimethyl fumarate for RRMS. Full details of the strategies used in each section were reported in the appendices or in the clarifications provided.

The search strategy for cost effectiveness studies (manufacturer's submission section 7.1.1 and7.1.2 P164 to 170) had some minor weaknesses but overall appears to be appropriate. A detailed commentary on the individual searches is provided in the Appendix 10.1.4.

The review of the literature did not appear to inform the manufacturer's choice of model. The choice of model was based on previously validated models produced for NICE and used in previous NICE technology appraisals for drugs for MS.[5-7] The model was first developed in TA32.[5] The models which were identified in the review generally adopted a similar premise of patients transitioning between EDSS health states to reflect disease and disability progression. This also forms the basis for the manufacturer's model as described in section 5.2.4. Only one model identified did not include EDSS states. This model estimated the costs and benefits associated with relapse and wheelchair dependence.

5.3 ERG's summary and critique of manufacturer's submitted economic evaluation

The evaluation conducted by the manufacturer combines clinical and economic data to evaluate the cost-effectiveness of dimethyl fumarate for the treatment of RRMS patients. The remainder of this section provides a summary and critique of the *de novo* model presented in the manufacturer's submission. A summary of the NICE reference checklist with the ERG's comments on whether the manufacturer's *de-novo* model has been judged to fulfil the NICE reference case is presented in Table 22.

In addition to the trial data discussed in section 4, the manufacturer utilised two exiting data sets to help inform parameters for the model, the UK MS survey and the London Ontario data. A brief summary of each of these datasets is presented with signposting to parameters which they inform and the section in which those parameters are discussed.

UK MS Survey

Data were collected by postal survey in February 2005. A total of 12,968 questionnaires were sent to people in the MS trust database; 2,508 were returned, of which 460 were censored and not used in the

evaluation; 2,048 of responses returned included evaluable information. The majority of the population (59.6%) were in EDSS states 4 to 6.5 and the mean age at diagnosis was 38.8 years. Of the 2,048 almost 29% had experienced a relapse during the last 3 months. The population comprised patients with three forms of MS, RRMS (35.5%), SPMS (37.2%) and PPMS (27.3%). The data from the survey has been used in previous NICE submissions.[5-7] These analyses have informed several journal publications.[39, 40] In the manufacturer's submission these data have been used to derive the EDSS health state costs by means of a regression analysis of the resource data collected in the survey (see section 5.2.10); and to derive the relationship between SPMS, RRMS and relapse utilities across EDSS states (see section 5.2.9).

London Ontario Dataset

The London Multiple Sclerosis Clinic (London Health Sciences Centre, Canada), established in 1972, provides long-term care for patients with multiple sclerosis from its referral area of south-western Ontario. Clinic and database characteristics have been extensively outlined in several publications.[41-43] Several analyses of these data appear to have been undertaken. However, much of the data remains CIC making a critique of the alternative analyses impossible. The referenced analysis in the dimethyl fumarate submission is based on the longitudinal follow-up of 1,099 consecutive MS patients evaluated at the MS Clinic of University Hospital, London, Canada, between 1972 and 1984. Of the total population 65.8% were classified as RRMS. Disability status scores were recorded annually, or as close to annual as possible using the Disability Status Scale (DSS). This scale antedated the extended DSS (EDSS). The DDS scale was modified several times to more accurately reflect the levels of disabilities clinically observed and renamed the EDSS.[44] A key change was the introduction of increments of 0.5 onto the original 0 to 10 point scale. The mapping of these data from DSS to EDSS does not appear to have been published. The London Ontario data were used to supplement the dimethyl fumarate trial data when deriving transition probability matrices for movement between EDSS states (see Section 5.2.4.2).

5.3.1 The manufacturer's economic evaluation compared with the NICE reference case checklist

Table 22 provides a summary of the NICE reference checklist with the ERG's comments on whether the manufacturer's de-novo model has been judged to fulfil the NICE reference case.

Table 22: NICE reference checklist

Attribute	Reference Case	Included in MS	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Comparator(s)	Alternative therapies in the NHS, including those currently regarded as current best practice	YES	Meets the scope set by NICE, and includes alternative therapies recommended by NICE. However, best supportive care is not included. In addition, two of the comparators are licensed and recommended for sub-populations but have been evaluated for the whole RRMS population.
Perspective – costs	NHS and PSS	YES	
Perspective - benefits	All health effects on individuals	YES	The utilities of both the MS patient and caregiver disutilities were incorporated in the model
Time horizon	Sufficient to capture differences in costs and outcomes	YES	30 year time horizon appears sufficient
Synthesis of evidence on outcomes	Systematic review	YES	
Outcome measure	QALYs	YES	
Health states for QALY measurement	Described using a standardised and validated instrument	YES	EQ-5D
Benefit valuation	Time Trade Off or Standard Gamble	YES	
Source of preference data	Representative sample of the public	YES	
Discount rate	3.5% on costs and health benefits	YES	
Equity weighting	No special weighting	YES	
Sensitivity analysis	Probabilistic sensitivity analysis	YES	The sensitivity analysis undertaken included probabilistic sensitivity analysis although a number of parameters were not assigned distributions.

5.3.2 Population

The scope distinguishes between the general RRMS population and subgroups for which fingolimod and natalizumab have been licensed and recommended.[6, 7] The population in this economic analysis is the RRMS population included in the dimethyl fumarate trials. As discussed in section 4.2.2, the population in the dimethyl fumarate trials differs from the general RRMS population only in that the trials' inclusion criteria required patients to have had a relapse within the 12 months prior to the start of the trial. This population is the closest to the scope population for which evidence is available, so the ERG does not consider the difference in the population to be a significant factor. Moreover, given that the aim of treatment is to reduce relapse rates, it seems sensible that only those patients suffering from relapse be prescribed treatment.

No analyses were done for the subgroup populations for which fingolimod and natalizumab have been licensed and recommended. The specific recommendations for treatment for the different drugs are presented in Table 23.

Table 23: Drug recommendations

Drug	Recommendation
Rebif 22µg Rebif 44µg Avonex Betaferon Glatiramer acetate	Initial treatment options for all RRMS patients suffering relapses at various rates. Patients must have had two relapse in the previous two years
Natalizumab	Initial treatment option for patients with rapidly evolving severe (RES) RRMS patients. RES is defined as two or more disabling relapses in 1 year, and one or more gadolinium-enhancing lesions on MRI or a significant increase in T2 lesion load compared with a previous MRI .[6]
Fingolimod	Treatment option for highly active RRMS patients. These are RRMS patients with a high level of disease activity despite initial treatment. This is 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 9 T2-hyperintense lesions in MRI or at least one gadolinium-enhancing lesion..[7] A non-responder was also defined in TA254 as a patient with an unchanged or increased relapse rate or on-going severe relapses as compared with the previous year.[7]

In the economic model, the baseline RRMS population is defined by age, gender and the EDSS score. The EDSS scale is described in section 3.4.2. The population distribution in the dimethyl fumarate trials over the EDSS scale is presented in Table 24. The ERG is not aware of a publication providing the general RRMS population distribution across EDSS states without the requirement of a relapse

within the previous 12 months. The ERG explores the sensitivity of the results to the population distribution over EDSS states in section 6.

Table 24: Baseline population distribution across the RRMS EDSS states. Taken from Table 42 (P175) in manufacturer's submission

RRMS EDSS state	0	1	2	3	4	5	6	7	8	9
% population	5.05	8.52	34.08	22.94	20.64	8.65	0.13	-	-	-

5.3.3 Interventions and comparators

The intervention considered is an oral medication, with an anticipated licensed dose of 240mg dimethyl fumarate, twice daily. The comparators included Rebif 22µg, Rebif 44µg, Avonex, Betaferon and glatiramer acetate, natalizumab and fingolimod at licensed doses.

These comparators have all been evaluated in RCTs for a population comparable to that of the dimethyl fumarate trials. However, as discussed in section 5.2.2, the treatments have not all been recommended or licensed for the same population. The inclusion of all these comparators is consistent with the scope, but not consistent with treatment recommendations or licenses.

Although the scope excludes best supportive care as a comparator, best supportive care or no treatment has a significant impact on the cost-effectiveness results because the economic model allows for discontinuation of treatment. Discontinuation of a treatment may be due to adverse events, death, moving on to an EDSS state of 7 or higher, and by developing SPMS. After discontinuation the model assumes that patients receive no treatment. This means that in any treatment arm of the model, a proportion of the patients are on treatment and a proportion are on no treatment – which in clinical practice the ERG believes would be 'best supportive care' or an alternative DMT. Due to discontinuation the proportion of patients off-treatment increases over time. The discontinuation rates used in the manufacturer's submission are described and discussed in section 5.2.8. Whether this reflects clinical reality is an issue. It is more likely that those patients who stop treatment for adverse effects will move to another active treatment if an alternative is available with a different adverse event profile. Further, it is also possible that those patients who progress to SPMS will receive some form of active treatment which has not been included in the model due to limitations in data and scope.

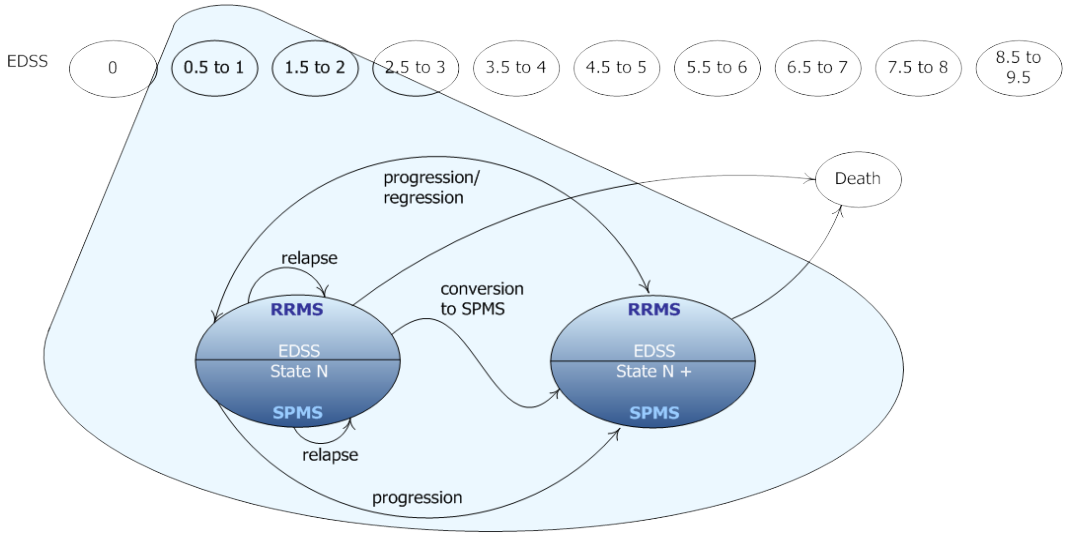
The results clearly indicate that had 'best supportive care' been included formally as a comparator, none of the DMTs being assessed would be considered cost-effective compared to NICE's conventional threshold for cost-effectiveness. When a patient discontinues treatment, they receive placebo or 'best supportive care' and make a quicker progression through EDSS states. The switch from treatment to no treatment leads to a significant reduction in treatment costs combined with a

small reduction in QALYs. This creates a situation where the ICER behaves in an unintuitive manner; that is, the more patients that discontinue treatment, the more cost-effective the drug appears. If 'best supportive care' were one of the comparators included in the scope, then the ICER of a drug would never be less than the ICER compared to 'best supportive care'.

5.3.4 Model structure

Within RRMS a patient may progress to a higher EDSS state or regress to a lower EDSS state, or progress to SPMS. Within SPMS a patient may progress to another EDSS state. Within an EDSS state a patient has an annual probability of relapse. This can be seen in Figure 1.

Figure 1: Model schematic. Taken from Figure 33 (P171) in manufacturer's submission



The manufacturer augmented an existing Markov model to model the natural history of patients with RRMS. The model incorporates disability progression, the progression from RRMS to SPMS, and the relapsing nature of the disease. The possible model transitions are represented in Figure 1 extracted from the manufacturer's submission. A patient with RRMS is considered to be in one of 10 EDSS states, 0 to 9. This is not a linear scale; the increase in disability moving from EDSS state 6 to 7 is greater than the increase in disability moving from EDSS state 3 to 4. As described in section 3.4.2, the EDSS scale comprises 20 states, increasing in increments of 0.5 scores, i.e. 5.0, 5.5, 6.0, 6.5, 7.0, 7.5. The rounding of EDSS states to only 10 states as done in the model may impact on the assessment of the cumulative probability of sustained progression of disability resulting in an overestimation of the rate of disability progression. A patient with RRMS may progress to a higher EDSS state or regress to a lower state. From any EDSS state a patient with RRMS may develop SPMS with an accompanying progression of 1 level in the EDSS scale.

Once the progression to SPMS has taken place a patient may not regress to RRMS, in addition only progression through the EDSS states is possible; no regression is allowed. The Markov model

incorporated annual transition probabilities between EDSS states for patients with either RRMS or SPMS, and between RRMS and SPMS. In any state, a patient may die. In any of the health states a patient may experience one or more relapses in any one year.

The model structure was based on a validated model developed by ScHARR for NICE.[5] This model has been used in three previous health technology submissions.[5-7] The submitted model structure differs from the original ScHARR model, TA32, in that there is the additional possibility for the RRMS population of regressing to lower EDSS states.[5] TA254 also did not allow regression to lower EDSS states.[7] However, TA127 did allow regression to lower EDSS states within its model.[6]

The inclusion of regression to lower EDSS states in the analysis reflects the actual experience of patients in the dimethyl fumarate trials and the experience of RRMS patients in the view of the ERG clinical expert. With its inclusion, it is the opinion of the ERG that the model predictions of the patients across the EDSS states seem reasonable compared to the distribution of dimethyl fumarate trial patients across the EDSS states within the time period of the trials. The disability progression outcome included in the analysis was 3 months sustained disability progression within 24 months, which is consistent with the approach in previous submissions.[5-7] According to the ERG clinical expert and the EMA guideline on clinical investigation of medicinal products for treatment of MS,[14, 45] confirmed sustained disability progression for 6 months may be more closely associated with permanent progression than 3 months sustained progression. The appropriate outcome to represent permanent progression is discussed in section 3.4.2. Given that 3 months sustained progression is not necessarily associated with permanent progression then, in the opinion of the ERG, the inclusion of regression to lower EDSS states in the model is reasonable. Excluding regression would result in the population progressing to higher EDSS states too quickly.

The drugs in the model affect the health of patients and cost to the health system through reduction in the annual relapse rate, the reduction in the annual risk of disability progression for patients with RRMS, and through the occurrence of adverse events. In the model, treatment has no effect on the probability of regression to lower EDSS states. When a treatment reduces the likelihood of disability progression, it increases the likelihood of remaining in the same disability state. No evidence was provided to support this assumption but the ERG considers it to be conservative with respect to dimethyl fumarate.

Discontinuation of the drug due to adverse events, death, moving on to an EDSS state of 7 or higher, is consistent with the Association of British Neurologist guidelines which suggest an EDSS of 6.5 as the maximum EDSS score for receipt of disease modifying treatment.[8]

Table 25: The sources for the natural history data

Type of transition	Source
Transitions between RRMS EDSS states	From EDSS states 0-7 for patients with RRMS to other states were derived from the placebo arms of the DF trials. From EDSS states 8/9 for patients with RRMS to other states were extrapolated using London Ontario data. All of these data were derived using the Multi State Markov statistical technique.
Transitions between SPMS EDSS states	Between SPMS EDSS states were based on the London Ontario dataset.
Transitions between RRMS and SPMS	From RRMS EDSS states to SPMS EDSS states were based on the London Ontario dataset.
Annualised relapse rate within each EDSS state	Based on the 12 months prior to enrolment in the DF trials for EDSS states 0-5. In the base case, the rates within states 6-9 were extrapolated from the lower states using data from the MS survey.
Mortality rate	Based on a publication by Pokorski.[34]

5.3.4.1 Baseline population

Patients entered the model with RRMS in one of the EDSS states according to the distribution presented in Table 24.

5.3.4.2 Transition probabilities

Transition probabilities were estimated for:

- Transitions between EDSS states for patients with RRMS;
- Transitions between RRMS to SPMS;
- Transitions between EDSS states for patients with SPMS.

A summary of the sources of the natural history data is presented in Table 25. Data for the transitions between EDSS states for patients with RRMS were obtained from two sources. Firstly, where the sample size was large enough, data were extracted from the placebo arms of the dimethyl fumarate trials. There were sufficient patients starting in EDSS states 0 to 7 in order to estimate a transition rate, therefore data on patient transitions from states 0 to 7 to other EDSS states were obtained from the trials. Secondly, data on transitions from RRMS EDSS states 8 to 9 to other RRMS EDSS states were obtained from the London Ontario dataset. The London Ontario dataset was briefly described in section 5.2. No other dataset was discussed as an alternative source. A multi-state Markov statistical (MSM) analysis was undertaken to estimate the transition probabilities. This estimates each transition probability simultaneously using a time to event distributional assumption, and ensures that the estimates are consistent with each other. The multi-state Markov statistical analysis appears appropriate.

As the dimethyl fumarate trial population was a general RRMS population, the transition probabilities between the SPMS EDSS states were based solely on the London Ontario dataset. No other dataset was discussed as an alternative source. Although the London Ontario dataset is quite dated and there is a lack of transparency surrounding the dataset, the ERG is not aware of an alternative source and so the ERG conducted some sensitivity analysis around these data in section 6.

In addition, the transition probabilities between RRMS and SPMS were also based solely on the London Ontario dataset. Trials of RRMS drugs are in general only two years in length so there is a lack of long-term data to inform these transition probabilities. The method of analysis of the transition probabilities between RRMS and SPMS was not reported. Since the transition to SPMS is independent of the treatment taken in the model, the effect of the transition on the ICER is similar to discontinuation due to adverse events. The earlier that patients transition to SPMS, the more cost-effective dimethyl fumarate.

Previous technology assessments have used a mixture of trial and London Ontario data for the RRMS EDSS state transitions as in the case of TA127 and only London Ontario data as in the case of TA254; both assessments used the London Ontario dataset for both the transition to SPMS from RRMS, and for the transitions between SPMS EDSS states.[6, 7]

5.3.4.3 Annualised relapse rate

Where the sample size was sufficiently high, the annualised relapse rate within each EDSS state was based on data from the 12 months prior to randomisation in the dimethyl fumarate trials. These were EDSS states 0 to 5. In the base case, the rates within states 6 to 9 were extrapolated from the lower states using data from the MS survey. The manufacturer conducted sensitivity analysis around these rates across all EDSS states (see section 5.2.13) by using estimates derived from the MS survey.

The UK MS survey emerged from the MS Risk Sharing Scheme, details of the population enrolled are discussed in section 5.2.

This approach was also used in two previous technology assessments (TA254 and TA127).[6, 7] The ERG is not aware of any alternative sources of data for relapse rates for EDSS states 6 to 9.

The annualised relapse rate estimates using the trial data and MS survey data, and using only MS survey data are reported in Table 26. The population of the dimethyl fumarate trials only differed from the general population stated in the scope by the requirement that a patient have had at least one relapse in the 12 months prior to the start of the trials. This may mean that the annualised relapse rates are higher than those in the general RRMS population. The ERG clinical advisor quoted 0.8 to be an appropriate annualised relapse rate in the general RRMS population. The MS survey used to extrapolate to EDSS states 6 to 9 in the base case analysis and to inform all of the EDSS states in the

sensitivity analysis has a broad population (35.5% RRMS; 37.2% SPMS; and 27.3% PPMS). Since the relapse rates in RRMS are higher than those in PPMS and many patients with SPMS may not have a relapse, the MS Survey estimates may be an underestimate of the annualised relapse rates.

Table 26: The annualised relapse rates for the RRMS EDSS states in the base case analysis using a combination of dimethyl fumarate trial and MS survey data, and using MS survey data only. Taken from Tables 47, 48 (P178/9) in manufacturer's submission

RRMS EDSS states	0	1	2	3	4	5	6	7	8	9
Base case (using DF trials)	1.26	1.32	1.32	1.35	1.36	1.43	1.18	1.23	1.23	1.23
MS survey estimates	0.71	0.73	0.68	0.72	0.71	0.59	0.49	0.51	0.51	0.51

5.3.4.4 Mortality rate

Patients with RRMS were assumed to have a higher mortality rate than the general population. The mortality multipliers for RRMS EDSS states that are multiplied with the general population mortality rates are presented in Table 27. The manufacturer referenced three papers supporting this assumption, but only used one of them, Pokorski et al.[34] In the manufacturer's response to the ERG's request for clarifications, it was stated that Pokorski et al was used in two previous submissions, which presented data under the following EDSS categories: 0 to 3.5, 4 to 7, and ≥7.5.[34] To derive the individual EDSS mortality rates from the grouped data, the manufacturer's analysis used linear interpolation, whereas an alternative method employed in TA254 fitted a curve to the data.[7] There does not appear to be a great difference between the approaches. The manufacturer conducted sensitivity analysis assuming a mortality rate the same as the general population. The ERG explores the impact of alternative mortality rates in section 6.

Table 27: The mortality multipliers for RRMS EDSS states compared to the mortality rate in the general population. Taken from Table 49 (P179) in manufacturer's submission

RRMS EDSS states	0	1	2	3	4	5	6	7	8	9
Mortality multiplier	1	1.3	1.6	1.68	1.76	1.84	2.71	3.57	4.44	5.31

5.3.5 Perspective, time horizon and discounting

The manufacturer's submission adopted an NHS and PSS perspective for the economic model, which is in accordance with the NICE scope. The time horizon was 30 years in the base case. As one example from the manufacturer's sensitivity analyses, the deterministic incremental cost-effectiveness ratios (ICERs) of dimethyl fumarate compared to Rebif 22µg with different time horizons modelled in sensitivity analysis in the manufacturer's submission are presented in Table 28. The ICER with a 30

year time horizon is close to that with a 50 year time horizon so it appears to be sufficiently long in order to capture the relevant differences in cost and clinical outcomes between the treatments. Discounting was appropriately conducted at a rate of 3.5%.

Table 28: The ICER of dimethyl fumarate versus Rebif 22 µg using the list prices for all drugs given different time horizons. Taken from Table 105 (P265) in manufacturer's submission

	Time horizon (Years)			
	10	20	30	50
ICER of dimethyl fumarate v Rebif 22µg with list prices	£293,292	£172,244	£142,283	£136,423

5.3.6 Treatment effectiveness and extrapolation

Two measures of treatment effectiveness were included in the model. These were the annualised relapse rate ratio and the relative risk of disability progression along the EDSS scale. Whilst not incorrect, it is often deemed more appropriate to use a hazard ratio for time to event outcomes rather that a relative risk. This point was raised by the ERG in the points for clarification and justification for the choice was requested. However, in addition to providing justification the manufacturer also presented an alternative analysis of the MTC using a hazard ratio of 3 months sustained progression as the outcome.

Both the annualised relapse rate ratio and the relative risk of progression have been used in previous NICE health technology appraisals.[5-7] In response to the ERG request for clarification on why a rate ratio was used the manufacturer stated that a risk ratio was chosen over a hazard ratio due to more consistent reporting of data in the trials for the MTC. The choice of a risk ratio over a hazard ratio maximised the number of trials included in the MTC. The manufacturer presented new analyses using what they described as a hazard ratio outcome. However, it appears that time to event data were not included in the analysis given the code provided and the outcome measure appeared to be a rate ratio instead. The ERG considers a risk ratio or a hazard ratio to be more appropriate than a rate ratio, so the rate ratio results are considered less reliable than the risk ratio results.

Confirmed disability progression sustained for three months at 24 months was chosen as the outcome measure for progression. The manufacturer's submission stated that this was preferred over confirmed sustained disability progression for 6 months at 24 months because more studies reported the 3 month outcome than the 6 month outcome and it resulted in a stronger network. Five of the seven comparators were included in the 6 month network compared to six comparators included in the 3 month network. Avonex was the one drug absent from the 3 months sustained disability network. Avonex was assumed to be the average of Rebif 22µg and Rebif 44µg. Sustained disability

progression for 6 months may be more closely associated with permanent progression than 3 months sustained progression.

The treatment effects were assumed to be the same across all EDSS states. The treatment effects over two years for all the treatments were derived from mixed treatment comparisons as described in section 4.4. The effectiveness of dimethyl fumarate compared to each comparator for both the annualised relapse rate ratio and the relative risk of progression is reported in Table 29. As discussed in Section 4.4, the mixed treatment comparisons were done in SAS. Overall, the MTC analyses seemed adequate given the data available. The calculation of risk ratios directly within an MTC rather than deriving them from odds ratios and baseline risks may produce slightly anomalous results.

Table 29: The treatment effect and adverse event profiles. The annualised relapse rate ratio (ARR) and risk ratio of progression (RRP) for dimethyl fumarate compared to seven comparators; and the difference in utility associated with adverse events. Taken from Figure 21 (P139); from Figure 28 (P145); adverse event utility from model in manufacturer's submission

	Dimethyl fumarate versus comparator		
	ARR†	RRP‡	Adverse event utility difference[]
Rebif 22µg	[redacted]	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]	[redacted]
Avonex	[redacted]	[redacted]	[redacted]
Betaferon	[redacted]	[redacted]	[redacted]
Rebif 44µg	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]
Natalizumab	[redacted]	[redacted]	[redacted]

†: ARR: Annualised relapse rate, 2 decimal places; ‡: RRP: Relative risk of 3 months sustained progression, 2 decimal places;][: A negative number indicates that the comparator has greater disutility than Dimethyl fumarate, 3 decimal places.

The time horizon of the model was 30 years. This compares with 30 years in TA32 and TA147 and 50 years in TA254.[5-7] Beyond 2 years a waning treatment effect was assumed for both the risk of progression and the annualised relapse rate ratio as presented in Table 30. Due to a lack of long term data, the waning effect was an assumption made by the manufacturer. The waning assumption was based on TA254.[7] There is therefore considerable uncertainty around this estimate and appropriate sensitivity analyses exploring alternative waning effects was conducted.

Table 30: The waning effect applied to all treatments in the years following the first two years of treatment. Taken from Table 53 (P183) in manufacturer's submission

Year	1	2	3	4	5	6	7	8	9	10+
Treatment efficacy (Waning effect)	100%	100%	75%	75%	75%	50%	50%	50%	50%	50%

5.3.7 Adverse events

An adverse event was included in the model if either there was an incidence rate of at least 5% in any of the dimethyl fumarate treatment arms from the DEFINE and CONFIRM trials; it is a common dimethyl fumarate adverse event on label and extracted in the systematic review; or it was an adverse event occurring at an incidence rate of at least 3% higher in the total dimethyl fumarate group than in the placebo group.

The treatment-specific annual incidence of adverse events was calculated from the systematic review of RCTs to inform treatment effectiveness, using a weighted average across studies. For each included adverse event, the proportion of each event that was serious was calculated using the number of serious events reported in the systematic review. This appears reasonable. The difference in the annual utilities between dimethyl fumarate and each of the comparators that are related to adverse events is reported in Table 29.

A mixed treatment comparison was also conducted to derive relative risks of adverse events for the same events for which incidence rates were calculated from the trials included in the systematic review. The manufacturer did not explain why incidence rates were calculated separately from the MTC and used in the economic model in place of the MTC results. The ERG will explore the use of the MTC results in Section 6.

5.3.8 Treatment discontinuation

The economic model allows for discontinuation of a treatment due to adverse events, death, moving on to an EDSS state of 7 or higher, and through progression to SPMS. This is consistent with the Association of British Neurologist guidelines which suggest an EDSS of 6.5 as the maximum EDSS score for receipt of disease modifying treatment.

The annual discontinuation rates included in the model are presented in Table 31. The relative risks of discontinuation for each comparator compared to placebo were estimated using the MTC as described in section 4.4. The baseline discontinuation rate for placebo was estimated from the placebo arms of trials included in the systematic review.

Table 31: The annual discontinuation risk used in the model. Taken from Table 52 (P183) in manufacturer's submission

	Dimethyl fumarate	Fingolimod	Glatiramer acetate	Avonex	Betaferon	Natalizumab	Placebo	Rebif 22 µg	Rebif 44 µg
Annual discontinuation risk	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Only the absolute discontinuation values were entered in the model. No probability distributions were assigned to relative risks of discontinuation. Since discontinuation may have a significant effect on the ICER through patients incurring far less treatment costs but only losing a small amount of benefit (i.e. QALYs), the mean ICER calculated from a probabilistic sensitivity analysis including probability distributions for the relative risks of discontinuation may be significantly different to a mean ICER calculated without those distributions.

The ERG believes the estimates of the relative risks reflect the trial evidence base. It is unclear to the ERG if there would be any differences between trial protocol and clinical practice discontinuation rates for oral compared to injectable drugs. The ERG explores the effect of different discontinuation rates in section 6.

5.3.9 Health related quality of life

5.3.9.1 RRMS and SPMS EDSS state utilities with and without relapse

Utilities were derived for each EDSS state (0 to 9) for both the RRMS population and the SPMS population. Utility decrements were also derived for relapses.

Table 32: EQ-5D scores for each EDSS state derived from DEFINE and CONFIRM. Taken from Table 57 (P58) in manufacturer's submission

EDSS score	EQ-5D index score (mean)	EQ-5D index score (SD)	Observations
0	0.88	0.17	513
1	0.83	0.19	846
2	0.78	0.19	3241
3	0.69	0.22	2185
4	0.63	0.22	2104
5	0.54	0.24	826
6	0.46	0.28	387
7	0.34	0.33	109
8	0.002	0.46	18
9	-0.17	0.29	11

Utilities for the RRMS population EDSS states were derived from all arms of the DEFINE and CONFIRM trials, pooling observations for each EDSS state (0 to 9) and calculating the mean EQ-5D index score for each state. The utilities are presented in Table 32. This appears appropriate. Sensitivity analysis using utility estimates from the MS survey was conducted by the manufacturer, but as discussed below there are some differences between the MS survey population and also the general RRMS population. In the base case analysis, these utilities were assumed to apply to patients with RRMS with no relapse.

To derive the utility estimates for patients with RRMS with relapse and for patients with SPMS with or without relapse, the manufacturer used utilities estimated from the MS survey data. There was a utility decrement of 0.009 for a relapse versus no relapse for patients with either RRMS or SPMS across all of the EDSS states. There was a utility decrement of 0.044 for SPMS versus RRMS across all of the EDSS states. The utility estimates derived from the MS survey are presented in Table 33. These decrements were applied to the utilities for the RRMS condition with no relapse to derive the utilities for RRMS with relapse and for SPMS with and without relapse.

The manufacturer references two sources for the utility data in Table 59, P192, of the manufacturer's submission. It is not clear which is the actual source, but of the two sources only one was a journal publication. This was a study by Orme *et al.* who undertook a multivariate linear regression of data obtained from the UK MS survey.[39] The UK MS survey is briefly described in section 5.2.

Table 33: Utility scores reported in the UK MS survey. Taken from Table 59 (P192) in manufacturer's submission

Clinical presentation	Disease type	EDSS state									
		0	1	2	3	4	5	6	7	8	9
No relapse	RRMS	0.909	0.844	0.745	0.611	0.654	0.558	0.495	0.437	-0.007	-0.151
	SPMS	0.865	0.800	0.701	0.568	0.610	0.514	0.451	0.393	-0.051	-0.195
Relapse	RRMS	0.900	0.835	0.735	0.602	0.645	0.548	0.485	0.427	-0.016	-0.160
	SPMS	0.856	0.791	0.692	0.559	0.601	0.505	0.442	0.384	-0.060	-0.204

Abbreviations: EDSS, Expanded Disability Status Scale; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis.

A systematic review was undertaken to identify HRQol data relevant to the decision problem, with the main focus being the identification of EQ-5D health state utility values in line with the preferred NICE method. The database search strategy for measurement and valuation of health effects (manufacturer's submission section 7.4.5 and section 7.4.6 P193 to198) utilised appropriate search terms but combined them in a manner that could result in relevant studies not being identified. However, additional searches of a number of resources were carried out and these may compensate for any deficiencies in the bibliographic database searching. Further details are provided in the Appendix, section 10.1.5.

The results were presented in a tabular format and difference between baseline EQ-5D scores and EDSS state distributions between studies identified in the review and DEFINE and CONFIRM were discussed. None of the studies identified provided utility data in a manner which allowed for its use in the manufacturer's model.

5.3.9.2 Adverse event disutilities

As previously reported in section 5.2.7, adverse events were considered for inclusion in the model if they had been reported in dimethyl fumarate studies; had an incidence rate of at least 5% in any of the treatment arms; were common dimethyl fumarate adverse events on label and extracted in the systematic review; or occurred at an incidence rate of at least 3% higher in the total dimethyl fumarate group than in the placebo group. The likely duration of these adverse effects were derived from the literature or expert option. The HRQoL and duration associated with each of the adverse events was then combined to derive disutility estimates. The final disutility assigned to each adverse event was based on both published sources and key opinion leaders (KOL) assumptions. All of the values were reported to have been validated by KOL, it would appear that the validation of utilities was undertaken by one clinical expert. The derivation of disutilities in this manner makes validation by the ERG difficult, although where possible calculations have been checked. It is the opinion of the ERG that the estimates used appear plausible, with the exception of influenza and flu-like symptoms. For these disutilities it is not clear to the ERG that the method of calculation is appropriate and an alternative calculation/analysis is presented in Section 6.2.5.2. Flushing with a high incidence rate is also assumed to have zero disutility. It is not clear if this is appropriate, although the ERG were unable to derive an alternative validated estimate.

5.3.9.3 Caregiver disutilities

Disutilities associated with caregivers were included in the model and these were estimated from the MS survey data. There is no published source that can allow these estimates to be evaluated, but the estimates are the same as in previous technology appraisals.[6, 7]

5.3.10 Resources and costs

The manufacturer's submission considers the resources necessary for the management of MS, as well as the resources required to provide DMTs. The included costs were presented in five categories:

- drug treatment costs;
- treatment administration costs;
- treatment monitoring costs;
- health state costs; and
- treatment related adverse event costs.

The cost analysis was undertaken taken from an NHS and PSS perspective. It appears that 2011/2 was used as the price year as some costs were reported as being inflated from 2011 to 2012 prices; however, this is not clear given the different sources used.

The manufacturers state that a systematic review of resource use in RRMS patients was not undertaken. This was justified by the manufacturer based on the availability of the UK MS Survey. As explained in section 5.2, the MS survey population is broader than the scope population. It has the following patient distribution across different forms of MS: 35.5% RRMS; 37.2% SPMS; and 27.3% PPMS. It is not clear that the resources accessed by this mixed population would be the same as those accessed by a general RRMS population. Also, as this survey is not publically available, it is not possible to assess how comprehensive this survey is in terms of resources used. Finally, as this survey was undertaken in 2005, there may be the potential for some relevant resource estimates to have been omitted.

The ERG believes that a systematic view was warranted, in particular, for the health state costs. Although the ERG did not conduct a systematic review of the literature, the ERG identified several publications not included in the manufacturer's submission,[46, 47] with Tyas et al (2007) also being based on the same UK MS survey used in the manufacturer's submission.[46] As will be described in more detail below, these publications provide varying estimates which could have been included in a sensitivity analysis in the manufacturer's submission.

5.2.10.1 Drug acquisition costs (Intervention and comparators)

The drug acquisition costs for dimethyl fumarate and its comparators are presented in Table 34. These costs were estimated using the doses per year in the drug licenses and the unit cost per dose from the British National Formulary. These annual acquisition costs have been correctly derived.

Table 34: Drug Acquisition costs. Taken from Table 65 (P204) in manufacturer's submission

Treatment	Annual Acquisition Cost
Dimethyl fumarate	£17,900
Fingolimod	£19,176
Glatiramer acetate	£6,841
Avonex	£8,531
Natalizumab	£14,690
Betaferon	£7,265
Rebif 22µg	£8,149
Rebif 44µg	£10,608

In the base case analysis, the list prices for all the drugs were used. In addition, the manufacturer included a sensitivity analysis where the dimethyl fumarate list price is replaced with manufacturer's proposed PAS price and compared with the list price for all the other drugs. The manufacturer's

proposed PAS price was stated as [redacted]. It is it important to note that this PAS priced has not been confirmed.

In addition the following sensitivity analyses were undertaken:

- The acquisition cost for fingolimod was varied from 20 to 50% less than its list price in 5% decrements;
- The cost of fingolimod was reduced by 35%.

These sensitivity analyses seem appropriate given that fingolimod is provided through a PAS and the cost reduction is not in the public domain.

Although the annual acquisition costs have been correctly derived, it should be noted that Avonex, Betaferon, Rebif 22µg , Rebif 44µg and glatiramer acetate are all available via an outcome based risk sharing scheme. The prices which were agreed as part of that scheme are publically available through a 2002 Department of Health circular.[13] The ERG has noted that the fingolimod submission also utilised the risk sharing scheme agreed prices. As these represent the price paid by the NHS, the ERG explores the scenario where the risk sharing prices for existing drugs are used along with the manufacturer's proposed PAS price for dimethyl fumarate in Section 6.

5.3.10.1 Administration and monitoring costs

Administration costs include any costs associated with a patient either self-administering or being administered a treatment. Dimethyl fumarate is an orally administered treatment and so, along with fingolimod, it does not incur administration costs. The injectable treatments incur administration costs and natalizumab requires a day case admission for administration, as outlined in Table 35.

Table 35: Administration costs. Taken from Table 66 (P204) in manufacturer's submission

Treatment	Annual administration cost (first year)	Resource use	Annual administration cost (second and subsequent years)	Resource use
Avonex Rebif 22µg Rebif 44µg Betaferon Glatiramer acetate	£99.00	3 hours of nurse's time to teach self-administration	£0	None
Dimethyl fumarate Fingolimod	£0	None	£0	None
Natalizumab	£6,224.00	Day admission	£6,224.40	Day admission

The administration costs were derived from the Department of Health's reference costs and appear appropriate.[36] The ERG clinical expert also felt that the resource use estimates for administration seemed appropriate. The natalizumab administration cost is based on the reference cost for "Medical

care of Patients with MS".[36] The administration costs included in the manufacturer's submission seem appropriate. However, it should be noted that TA127 indicated that natalizumab has an annual administration cost of £1,062 (2008 price year), which is substantially lower.[6]

Monitoring costs are the costs associated with any additional tests necessary for patients whilst on treatment. The costs associated with monitoring for each treatment are presented in Table 36. In contrast with other DMTs, dimethyl fumarate requires annual renal function tests. The submission indicates that the unit costs were derived from the 2011/12 Department of Health reference costs; however, the reference provided is for the 2010/11 reference costs. The ERG have checked this and it would appear that the 2011/12 prices have been used.[36] A weighted average was used for the neurology visit and MRI scan costs. It is not clear how these weighted averages were calculated, but the unit costs used in the submission have only small differences in price compared with the average unit costs in the Department of Health reference costs and so it is unlikely to affect the cost-effectiveness results.

Table 36: Annual cost and resource use of monitoring for each treatment. Taken from Table 67 (P205) in manufacturer's submission, and annual resource use estimates of the manufacturer and the ERG clinical expert

Treatment	Resource use in Year 1 (manufacturer's submission)	Cost in Year 1 (manufacturer's submission)	Resource use in year 1 (adjusted following clinical expert discussion)	Resource Use in subsequent years (manufacturer's submission)	Cost in subsequent years	Resource use in subsequent years (adjusted following clinical expert discussion)
Avonex Rebif 22µg Rebif 44µg Betaferon	3 neurology visits 3 full blood counts 3 liver function tests	£1,776.86	2 neurology visits (baseline and 12 months) 3 full blood counts 3 liver function tests 4 MS nurse visits (at 1, 3, 6 and 12 months)	1 neurology visit 2 full blood counts 2 liver function tests	£594.75	1 neurology visits 2 full blood counts 2 liver function tests 2 MS Nurse visits
Dimethyl Fumarate	3 neurology visits 3 full blood counts 3 liver function tests 3 renal function tests	£1,780.55	3 neurology visits 3 full blood counts 3 liver function tests 3 renal function tests 4 MS nurse visits	1 neurology visit 2 full blood counts 2 liver function tests 2 renal function tests	£597.21	2 neurology visits 2 full blood counts 2 liver function tests 2 renal function tests 2 MS nurse visits
Fingolimod	3 neurology visits 3 full blood counts 3 liver function tests 1 basic	£2,431.09	3 neurology visits 3 full blood counts 3 liver function tests 1 basic metabolism test 1 ophthalmology	1 neurology visit 2 full blood counts 2 liver function	£597.21	2 neurology visits 2 full blood counts 2 liver

	metabolism test 1 ophthalmology visit 1 patient observation after first administration by healthcare professional		visit 1 patient observation after first administration by healthcare professional 4 MS nurse visits	tests 2 basic metabolism test		function tests 2 basic metabolism test 2 MS nurse visits
Glatiramer acetate	2 neurology visits 2 full blood counts 2 liver function tests	£1,184.57	2 neurology visits 2 full blood counts 2 liver function tests 4 MS nurse visits	1 neurology visit	£589.83	1 neurology visit 2 MS nurse visits
Natalizumab	2 neurology visits 2 full blood counts 2 liver function tests 1 MRI scan	£1,334.07	3 neurology visits 2 full blood counts 2 liver function tests 1 MRI scan 4 MS nurse visits	2 neurology visits	£1,179.66	2 neurology visits 1 MRI scan 2 MS nurse visits

The ERG clinical expert considered the resource use assumptions presented in Table 36. It was confirmed that, as per the manufacturer's submission, a day case admission for patient observation following the initial administration of fingolimod was standard practice and an appropriate inclusion in the resource use of fingolimod in year 1, and is in line with TA254.[7] However, the manufacturer's estimated neurology visits for Avonex, Rebif 22µg, Rebif 44µg, Betaferon in year 1 were considered too high; the estimated neurology visits for natalizumab in year 1 were too low; the estimated neurology visits for dimethyl fumarate and fingolimod in subsequent years were too low; and finally that MRI scans are required annually for patients taking natalizumab.

The ERG clinical expert also highlighted the exclusion of resource use estimates for MS nurse visits. In their opinion, it would be expected that MS patients on injectable disease modifying treatments would be assessed by an MS nurse four times in the first year and twice annually in subsequent years. MS nurse visits have not been included in the manufacturer's submission and it is not clear why this resource use was omitted.

The assumption that additional blood tests, such as liver function tests and full blood counts, will not require any additional visit costs was also considered to be appropriate.

Therefore table 36 presents alternative resource use estimates that the ERG believes may better reflect current practice in the UK. The ERG evaluates the effect of these alternative assumptions in section 6.

5.3.10.2 Health state costs

The health state costs include the cost associated with each EDSS state, the average cost of a relapse and the cost of adverse events.

Relapse costs

The resource use estimates for EDSS state costs and relapse costs were derived from the UK MS Survey. Both the EDSS state costs and the average cost of relapse were estimated using seemingly unrelated regressions (SUR). As explained in the manufacturer's submission, the SUR accounts for correlation between costs, allowing for both dependence in costs within a patient and independence between patients. Using SUR to estimate the EDSS state costs has been undertaken in a previous submission (TA127); however, very little detail was provided on how these costs were estimated, in either the manufacturer's submission or in TA127.[6] It is therefore not possible to fully assess whether these estimates were appropriately derived.

Using SUR, the average cost of a relapse was estimated to be £2,028 for both RRMS and SPMS. The ERG clinical expert estimated that not more than 20% of relapses would result in inpatient admissions, with the majority of relapses being treated at relapse clinics and a proportion of this group admitted as day case patients for IV steroids. An estimate for non-elective admitted patient care and outpatient procedure used in TA254 is £3,039.[7] The cost for the other 80% may be very low, so £2,028 appears to be a high estimate, or it appears to be assumed that a higher proportion of patients are admitted to hospital.

The ERG has identified 4 different estimates for the cost of a relapse, including the estimate used in this submission. Three of these, including this submission, are estimates reported to have been derived from the same MS survey data using the same SUR analysis method. These are presented in Table 37.

Table 37: Different relapse cost estimates and their sources

Reference	Relapse cost estimate (£)	Price year	Data source
Dimethyl fumarate submission	2028	2012	MS survey
TA127 (natalizumab)	228	2005	MS survey
Tyas et al	1623	2007	MS survey
TA254 (fingolimod)	3039	2010	NHS Reference costs

As can be seen from Table 37, the three estimates derived from the MS survey vary significantly. A possible explanation for this may be varying unit costs. For example, in the NHS Reference costs, depending on the reference code used, the unit cost of an NHS trust, consultant led, non-admitted, face to face, multi-professional, follow up attendance varied from £215 to £647. However, this may be just one explanation. As the ERG cannot fully evaluate either the MS survey resource use estimates or the SURs used to calculate the costs, it is unclear why the costs vary so substantially.

The fingolimod submission (TA254) estimate is based on a non-elective admitted patient care and outpatient procedure tariff.[7] This may not be appropriate given the ERG clinical expert estimate of not

more than 20% of relapses resulting in inpatient admissions. As the manufacturer's estimate is at the upper end of the estimates identified, and given the ERG clinical expert estimation, the ERG expects that the relapse costs in this submission have been overestimated. The ERG will explore alternative relapse cost assumptions in section 6.

EDSS state costs

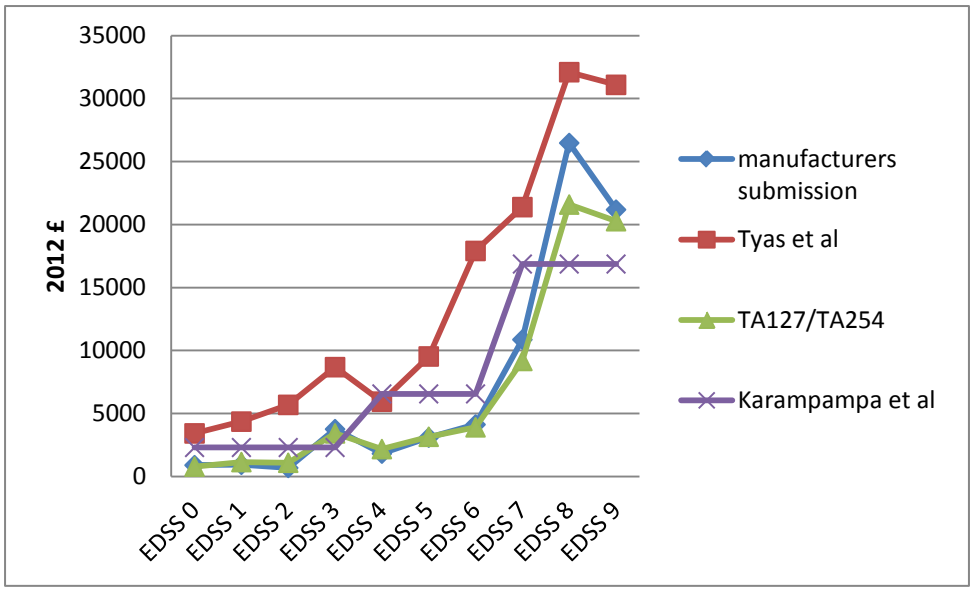
The annual cost, provided in the manufacturer's submission, for each EDSS state is presented in Table 38. The submission provides different EDSS state costs for patients with RRMS compared with SPMS patients, which seems appropriate and is in line with previous technology appraisals (this was not done in the fingolimod assessment (TA254) and this was heavily criticised; an additional cost for SPMS patients was included in TA127).[6, 7]

Table 38: EDSS state costs in the model (£). Taken from Table 70 in manufacturer's submission

Disease Type	EDSS 0	EDSS 1	EDSS 2	EDSS 3	EDSS 4	EDSS 5	EDSS 6	EDSS 7	EDSS 8	EDSS 9
RRMS	903	939	688	3,765	1,824	3,094	4,130	10,871	26,478	21,187
SPMS	1,217	1,254	1,002	4,079	2,138	3,409	4,444	11,185	26,793	21,502

The ERG has identified four different estimates for EDSS state costs. These estimates are compared in Figure 2. Three of these estimates are based on the MS survey including this submission.[6, 7, 46] The fourth estimate (Karampampa et al.) was based on a UK sample of 194 patients, 72% of which were RRMS patients and 75% of whom were receiving DMTs.[47]

Figure 2: Alternative EDSS state costs for RRMS patients



Note: Karampampa et al present EDSS state costs grouped in three categories: 0 to3, 3 to 6.5 and 7 to 9; the average costs were calculated across the states to estimate 10 values.

For this submission, TA127 and Tyas et al, the same resource data (the UK MS survey) was used, with different unit costs attached.[6, 46] These unit costs varied from using the PSSRU, NHS reference costs and a combination of both. A SUR was then run for each to obtain the estimates seen in Figure 2. These variations in unit costs may explain the variation in estimates seen in Figure 2. Karampampa et al used both the PSSRU and the NHS reference costs but used different resource use estimates.[47]

While the ERG expects that the differing unit costs used to estimate the EDSS states costs may explain some of the variations, it is not clear if this is the only factor that could explain the variation. In Tyas et al, medical and non-medical costs are distinguished but no such distinction is explicitly made in other analyses. It is possible that some non-medical costs are not relevant to the NHS PSS perspective. It is unclear which set of estimates is the most appropriate. It can been seen in Figure 2 above that there are different distributions of EDSS state costs depending on the source used, which may also effect the ICER. It may be expected that the higher the costs associated with EDSS states the greater the benefit from a reduction in disability progression and the higher the ICER of dimethyl fumarate. If costs were lower in low EDSS states and higher in high EDSS states, the effect is less predictable, but this may result in a higher ICER for a more effective drug such as dimethyl fumarate because there are less immediate benefits from a reduction in disability progression.

The ERG will explore alternative EDSS state costs, and their effect on the ICERs, in section 6.

Costs of adverse events

The cost of treatment related to adverse events incurred by patients is provided in the manufacturer's submission for serious and non-serious adverse events associated with dimethyl fumarate and its comparators. These are presented in Table 39.

As previously reported in section 5.2.7 and section 5.2.9.2, adverse events were considered for inclusion in the model under certain criteria and the incidence of these events were obtained from the trials identified in the systematic review. The resource use costs were sourced from PSSRU 2011, NHS reference costs 2011-12, as well as some assumptions made by the manufacturer.[36, 48] The adverse events list appears to be a fairly exhaustive list and the costs were validated by clinical expert opinion. However, flushing has a high incidence and is mostly associated with dimethyl fumarate and is assumed to incur no cost. It is not clear if this is appropriate, but the ERG does not have an alternative estimate.

Table 39: Costs of treatment for Adverse Events. Taken from Table 72 in manufacturer's submission

Adverse Event	Cost per event (Non-Serious, £)	Cost per serious event (£)
Abdominal pain	53.00	53.00
Abdominal pain upper	53.00	53.00
Alanine transaminase increased	0.00	0.00
Arthralgia	53.00	53.00
Atrioventricular conduction block	469.63	1,833.54
Back pain	53.00	53.00
Bradycardia	1,091.90	1,357.49
Chest pain	542.03	542.03
Cough	0.00	0.00
Depression	265.00	636.00
Diarrhoea	0.00	0.00
Fatigue	0.00	0.00
Flu-like symptoms	50.55	251.48
Flushing	0.00	0.00
Gastroenteritis	642.38	1,172.99
Headache	661.03	661.03
Influenza	50.55	502.97
Leukopenia	0.00	0.00
Lower respiratory tract infections	860.97	860.97
Nausea	0.00	53.00
Pain in extremity	125.50	125.50
Pruritus	0.00	53.00
Rash	564.22	564.22
Urinary tract infection	86.92	173.85

5.3.11 Discounting

The costs and utilities were discounted at an annual rate of 3.5%, with is in line with the NICE reference case. Within the sensitivity analysis, the discount rates were varied as follows:

- Costs and benefits discounted at 0%
- Costs and benefits discounted at 6%
- Costs discounted at 0%, benefits at 6%
- Costs discounted at 6%, benefits at 0%
- Costs discounted at 1.5%, benefits at 3.5%
- Costs discounted at 3.5%, benefits at 1.5%

5.3.12 Cost effectiveness results

This section presents the manufacturer's base case cost-effectiveness results, and the results of their sensitivity and scenario analyses.

5.3.12.1 Base case results (list prices for all drugs)

The base case analysis was based on the list prices for all treatments. The deterministic full incremental cost-effectiveness results, which exclude dominated treatments, are presented in Table 40. The manufacturer did not provide mean ICER results calculated from probabilistic sensitivity analysis (PSA). In the manufacturer's response to clarifications, it was claimed that the results from probabilistic sensitivity analysis were not stable. This was based on an adapted model with additional probability distributions. The ERG presents PSA results in section 6.

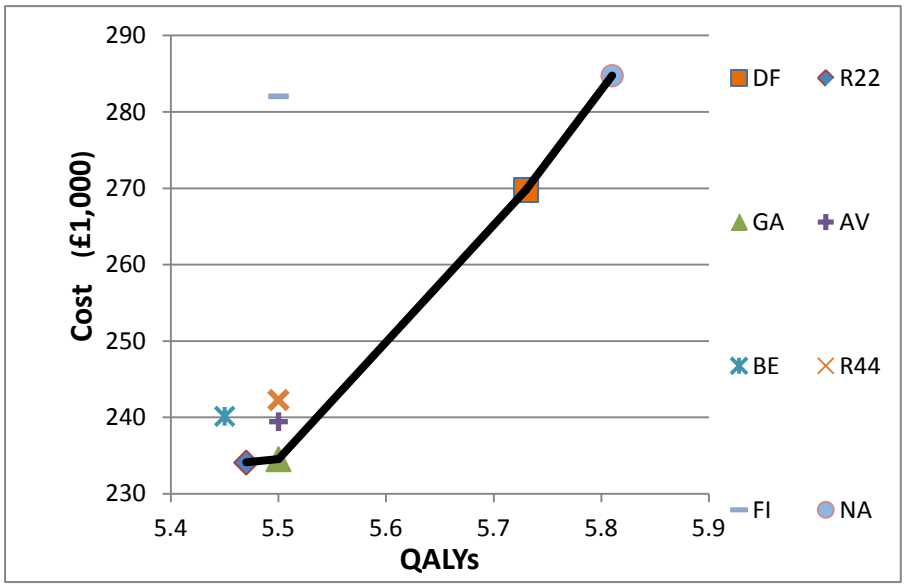
In Table 40 the incremental cost-effectiveness ratios of each treatment compared to the next most costly alternative is calculated. For example, the incremental cost-effectiveness ratio (ICER) of natalizumab is £173,745, which is compared to dimethyl fumarate, the next most costly treatment. A dominated treatment is more costly and less effective when compared to another treatment (strictly dominated) or when compared to a combination of two other treatments (dominated by extension). The ICER of dimethyl fumarate is £159,295 per QALY.

Table 40: The deterministic incremental cost-effectiveness results using the list prices for all the drugs. Taken from Table 92 (P238) in manufacturer's submission

Treatment	Total cost	Total QALYs	Incremental costs	Incremental QALYs	ICER
Rebif 22µg	£234,103	5.47	-	-	
Glatiramer acetate	£234,547	5.50	£445	0.03	£15,026
Dimethyl fumarate	£269,798	5.73	£35,250	0.22	£159,295
Natalizumab	£284,763	5.81	£14,965	0.09	(£173,745)

The cost and QALY results for each treatment are shown in Figure 3. The cost-effectiveness frontier is also presented. All treatments not on the frontier are dominated by another treatment or a combination of two other treatments. The only drug that produced more QALYs than dimethyl fumarate was natalizumab. The ICERs for pair-wise comparisons were as follows. Natalizumab was not cost-effective compared to dimethyl fumarate with a cost-effectiveness ratio of £173,745. The only treatment that was dominated by dimethyl fumarate, i.e. was more costly and less effective, was fingolimod. Dimethyl fumarate was not cost-effective compared to the other comparators with the ICER for dimethyl fumarate ranging from £106,127 to £159,295.

Figure 3: The cost-effectiveness plane showing the deterministic costs and QALYs of each treatment, and the cost-effectiveness frontier. All treatments not on the frontier are dominated by one or a combination of other treatments. The list prices are used for all drugs.



5.3.12.2 PAS price for Dimethyl fumarate

In a sensitivity analysis the manufacturer presents incremental cost-effectiveness results using the manufacturer's proposed PAS price for dimethyl fumarate and the list price for all other treatments. The list price for dimethyl fumarate is £17,900 and the manufacturer's proposed PAS price is [redacted]. The full incremental cost-effectiveness results excluding the dominated treatments are presented in Table 41. The cost and effectiveness results for each treatment along with the cost-effectiveness frontier are presented in Figure 4. Due to the reduced price of dimethyl fumarate while maintaining the list price for all other drugs, the ICER of dimethyl fumarate has been reduced from £159,295 per QALY to £19,057 per QALY.

For pair-wise comparisons, the ICER of natalizumab compared to dimethyl fumarate increased to £534,047. Dimethyl fumarate now dominated, i.e. was less costly and more effective, Rebif 44µg, Fingolimod, Betaferon and Avonex. The ICER of Dimethyl fumarate was £18,581 compared to Rebif 22µg.

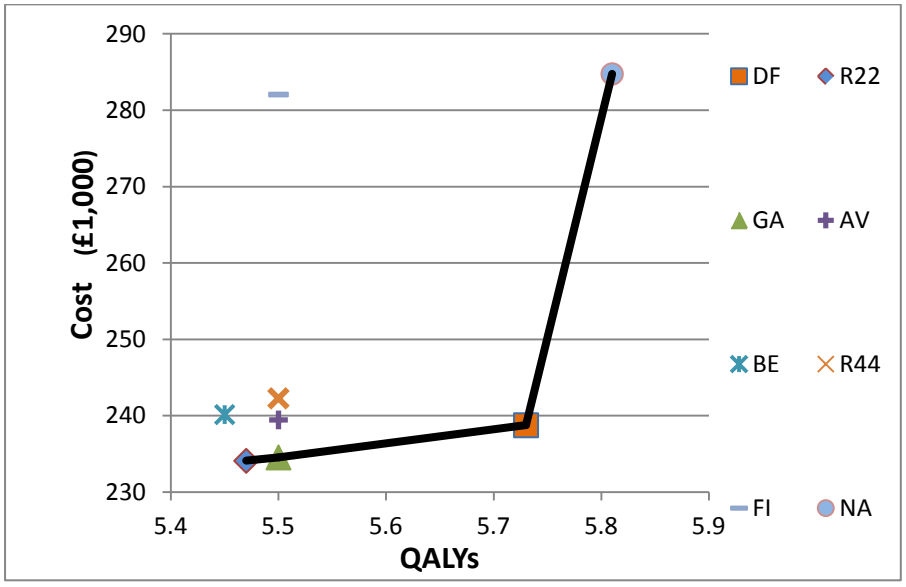
The actual price paid for other drugs including fingolimod, Avonex, Betaferon, Rebif 22µg, Rebif 44µg and glatiramer acetate by the NHS in England and Wales is less than the list price as discussed in Section 5.2.10.1. The ERG considers that a more appropriate analysis would be to evaluate the cost-effectiveness of dimethyl fumarate utilising discounted prices for every drug where possible. This is explored in Section 6.

Table 41: The deterministic incremental cost-effectiveness results using the manufacturer's proposed PAS price and the list price for all other drugs. Taken from Table 93 (P238) in manufacturer's submission

Treatment	Total cost	Total QALYs	Incremental costs	Incremental QALYs	ICER
Rebif 22µg	[redacted]	[redacted]	[redacted]	[redacted]	
Glatiramer acetate	[redacted]	[redacted]	[redacted]	[redacted]	£15,026
Dimethyl fumarate	[redacted]	[redacted]	[redacted]	[redacted]	£19,057
Natalizumab	[redacted]	[redacted]	[redacted]	[redacted]	(£534,047)

Brackets indicate that the ICER is for the reverse comparison of treatments than that stated, i.e. natalizumab versus dimethyl fumarate.

Figure 4: The cost-effectiveness plane showing the deterministic costs and QALYs of each intervention, and the cost-effectiveness frontier. All treatments not on the frontier are dominated by one or a combination of other treatments. The manufacturer's proposed PAS price for dimethyl fumarate is used and list prices are used for all other drugs.



5.3.12.3 PAS price for fingolimod as well as dimethyl fumarate

It is known that there is a PAS price to the NHS for fingolimod, but this price is not publically available. The manufacturer therefore conducted a threshold analysis reducing the list price of fingolimod in 5% increments until dimethyl fumarate no longer dominated fingolimod (i.e. was no longer cheaper and more effective). Dimethyl fumarate was still cheaper as well as more effective than fingolimod until the price of fingolimod was reduced by more than 55% from the list price, reducing the price of fingolimod from £19,176 to less than £8,629. The cost-effectiveness ratio of dimethyl fumarate compared to fingolimod was still less than the cost-effectiveness threshold of £30,000 until the price of fingolimod fell to less than £6,712.

5.3.12.4 Alternative rate ratio analysis

In the response to the points of clarifications raised by the ERG, the manufacturer provided incremental cost-effectiveness results using the hazard ratio instead of the risk ratio for 3 months sustained progression. However, it appeared to be a rate ratio rather than a hazard ratio given the code provided. This is discussed in section 4.4. The ERG considers the rate ratio results to be less reliable than the risk ratio results provided in the original submission.

5.3.12.5 Probabilistic sensitivity analysis

Probabilistic sensitivity analyses (PSA) were performed and results reported for both scenarios where the list price and the manufacturer's proposed PAS price were used for dimethyl fumarate.

Probability distributions were specified for most parameters, but they were not specified for

- The adverse event estimates;
- The treatment waning effect after 2 years;
- The annual discontinuation risk;
- The baseline distribution of the population across EDSS states.

The treatment waning effect and the annual discontinuation risk have a significant effect on the ICER of dimethyl fumarate because disease progression is the driver of the model and patients that discontinue treatment receive best supportive care which is more cost-effective than being on treatment.

In addition, fixed treatment effects were assumed in the MTC analyses given the lack of trials to estimate a between-study estimate, so it is possible that the uncertainty in the treatment effect estimates is underestimated. If the uncertainty in the treatment effect is underestimated then the uncertainty in the cost-effectiveness of dimethyl fumarate will also be underestimated. The distributions specified for the utility estimates and the natural history of relapse rates were assumed to have a standard error of 10% of the mean.

The PSAs were run using 1,000 iterations. The outputs of the PSA were:

- Mean cost and QALY estimates for each treatment along with 95% confidence intervals;
- Scatter plots of the results for the iterations for dimethyl fumarate compared to each comparator;
- Cost-effectiveness acceptability curves for dimethyl fumarate compared to each comparator.

Mean ICERs for dimethyl fumarate were reported for the PSA. In the response to the points for clarification raised by the ERG, the manufacturer stated that the mean ICER PSA results were not stable with 10,000 iterations. The ERG will explore this in section 6.

The following are the results presented in the original submission.

- When the list price for dimethyl fumarate was used for all drugs, the probability that dimethyl fumarate was cost-effective compared to all comparators apart from fingolimod was close to zero at a threshold of £30,000 per QALY. When compared to fingolimod, the probability that dimethyl fumarate was cost-effective was close to 1.
- When the manufacturer's proposed PAS price was used for dimethyl fumarate and the list prices for all other drugs, the probability that dimethyl fumarate was cost-effective was 1 when compared to fingolimod and natalizumab, and roughly 0.83 when compared to Rebif 44µg. It was close to 0.75 when compared to Avonex, Betaferon, and it was close to 0.5 when compared to glatiramer acetate and Rebif 44µg.

The ERG feels that the use of list prices for all drugs is an acceptable base case analysis, but that an analysis that includes discounted prices for all drugs is more useful to the NHS. An analysis comparing the proposed PAS price for dimethyl fumarate with the list prices for all other drugs is the least appropriate.

5.3.12.6 One-way sensitivity analyses

The manufacturer conducted several one-way analyses assuming the list price for dimethyl fumarate and assuming the manufacturer's proposed PAS price for each of them. The ICER was calculated for dimethyl fumarate compared to each comparator. Each parameter tested was varied by +/-20%. It is not clear that this range is adequate in every case, and the ERG will explore alternative ranges in Section 6.

List price analyses

When the list price for dimethyl fumarate was used, the results did not change across the full range of values tested for each parameter.

PAS price analyses

When the manufacturer's proposed PAS price was used while using the list prices for all other drugs, dimethyl fumarate remained cost-effective or not cost-effective against each comparator across the full range of values tested for each parameter when compared to a cost-effectiveness threshold of £30,000 per QALY except for the following cases.

Dimethyl fumarate was no longer cost-effective when compared to glatiramer acetate, Avonex and Rebif 22µg when either the 3 months sustained progression rate for dimethyl fumarate was increased by 20% or the progression rate for the comparator was decreased by 20%. When compared to Betaferon, dimethyl fumarate was no longer cost-effective only when the comparator progression rate fell by 20%.

5.3.12.7 Two-way sensitivity analyses

The two-way sensitivity analyses were conducted on effect parameters relative to placebo. These were the relapse rate, disability progression risk ratio, and drop outs. In the model all effect parameters were relative to placebo. In these analyses, the effect parameter relative to placebo was either increased for both dimethyl fumarate and the comparator at the same time or decreased at the same time. The limitation of this approach is that the full range of plausible values of the effect parameters of dimethyl fumarate relative to the comparator is not explored. The ERG conducts additional sensitivity analyses around these parameters in Section 6.

In these analyses, the only cases where the cost-effectiveness of dimethyl fumarate changed relative to a threshold of £30,000 were compared to glatiramer acetate and to Rebif 22µg, when the manufacturer's proposed PAS price was used, and when the lower confidence interval of the risk ratios of disability progression was used for both dimethyl fumarate and the comparator compared to placebo at the same time. Dimethyl fumarate ceased to be cost-effective in these scenarios.

5.3.12.8 Scenario analyses

The manufacturer conducted several scenario analyses. A few of these may be considered one-way sensitivity analyses on model parameters. These made the following alternative assumptions:

- No treatment waning effect;
- 0 to 6% discount rates;
- 1 to 50 year time horizons;
- Mortality rate of MS population equal to the general population;
- Annualised relapse rates relative to placebo at 95% upper and lower confidence limits;
- Disability progression relative to placebo at 95% upper and lower confidence limits;
- London Ontario transition matrix for RRMS-RRMS transitions.

List price analyses

There was only one occasion when the cost-effectiveness conclusion for dimethyl fumarate changed relative to a threshold of £30,000. Dimethyl fumarate changed from being cost-effective compared to natalizumab to not cost-effective when upper 95% confidence interval for the disability progression risk ratio was used for dimethyl fumarate.

PAS price analyses

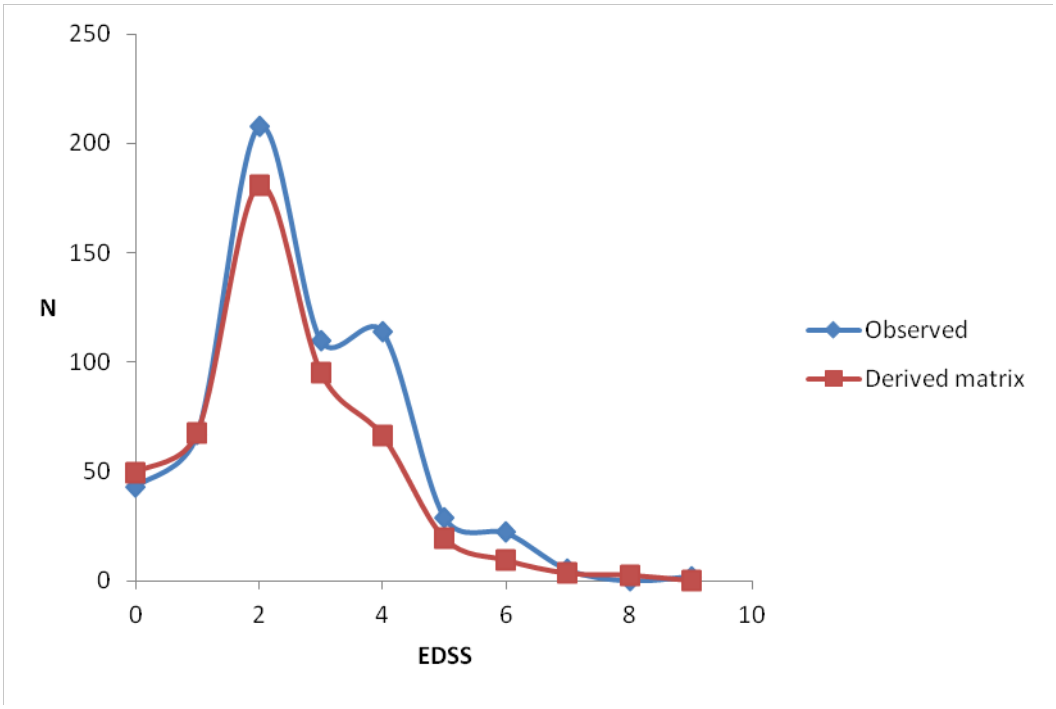
The only occasion when the cost-effectiveness conclusion for dimethyl fumarate changed relative to a threshold of £30,000 was when the time horizon was reduced to 10 years or less, which the ERG considers to be too short a time horizon. This applied when dimethyl fumarate was compared to Rebif 22µg, glatiramer acetate; and to Betaferon when the time horizon was 5 years or less.

5.3.13 Model validation and face validity check

The manufacturer conducted validation tests on the model by comparing the predicted distribution of the population across EDSS states from the model with that observed in the dimethyl fumarate trials for the placebo and dimethyl fumarate populations for the first and second years each. The predicted cumulative survival was also compared to the cumulative survival estimates reported in Kingwell et al.[48]

Figure 5 presents the predicted population distribution from the model compared with the actual distribution at the end of year 2 for the dimethyl fumarate population. It appears to be a reasonable fit, perhaps with a slightly higher proportion of patients in EDSS states 2 and 4 in the model output. This may slightly reduce progression over the long-run, which would favour dimethyl fumarate.

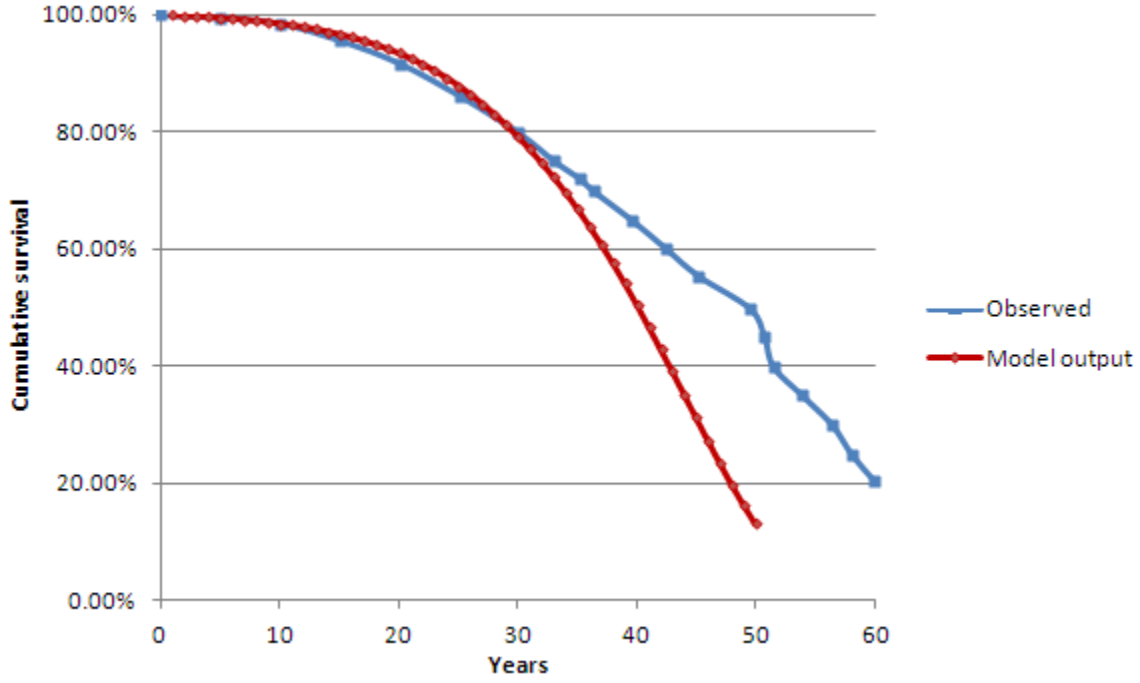
Figure 5: Line graph of model EDSS state occupancy at end of year 2 (derived matrix), and observed patient distributions at week 96 in the clinical trial for the dimethyl fumarate population. Taken from Figure 46 (P270) in manufacturer's submission



Abbreviations: EDSS, Expanded Disability Status Scale.
Mean square error = 358.41, Root mean square error = 18.93.

Figure 6 presents the predicted cumulative survival compared to the results reported in Kingwell et al.[48] The base case time horizon for the model was 30 years and the model prediction fits the Kingwell data well for that period.

Figure 6: Comparison of cumulative survival in the model against observed studies. Taken from Figure 47 (P271) in manufacturer's submission



Mean square error = 2.45%, Root mean square error = 15.6%

Mean square error until 0 - 35 years: 0.05%, 35 - 50 years: 4.42%

The manufacturer made several face validity checks inbuilt in the model. The ERG believes that the model is internally valid.

5.4 Conclusions of the cost effectiveness section

The *de novo* economic evaluation was reasonably well conducted and reported. The most significant factor in the economic analysis is the relative price of the drugs. When the list prices for all of the drugs are used, dimethyl fumarate is not cost-effective compared to most of the comparators included in the analysis. When the manufacturer's proposed PAS price is used for dimethyl fumarate and the list price is used for all other treatments, dimethyl fumarate becomes cost-effective. The ERG considers it more appropriate to compare reduced prices across all the treatments where there are reduced prices and these are known. The ERG explores this in section 6.

The driver of the model is disease progression through EDSS states. Across the sensitivity analyses conducted by the manufacturer, the results were mainly only sensitive to variation in the rate of 3 months sustained progression. The mixed treatment comparisons that were conducted to obtain the

relative treatment effects appeared to be appropriately conducted. There was considerable uncertainty around these estimates. Although a fixed treatment effect was the most practical assumption in most of the mixed treatment comparisons, the estimates of confidence intervals may be slightly underestimated as a result.

Although EDSS states may be imperfect at mapping the change in disease state of a patient, modelling the progression through EDSS states is the common approach to modelling the disease pathway in technology appraisals submitted to NICE. The model allowed patients to regress to lower EDSS states, which appeared to be appropriate given that patients did in fact regress to lower EDSS states in the dimethyl fumarate trials. The model was critically assessed using the Phillips checklist, which is reported in the Appendix 10.3.

All aspects of the economic evaluation were consistent with the NICE reference case with the qualification that best supportive care was excluded as a comparator. This, however, was consistent with the scope. The comparators and outcomes were also consistent with the scope. Given the data and assumptions included in the model, none of the treatments were cost-effective compared to placebo, and this had the counterintuitive effect that the higher the drop-out rate, the more cost-effective the treatment.

The principal scope population was the general RRMS population and the population for the treatment effectiveness estimates was for a population with a recent relapse. Although these populations differ slightly, the ERG does not consider this to be significant. In clinical practice, one or more recent relapses is used as an indication for existing DMTs, so the trials are representative of the appropriate clinical population. The modelled population does not distinguish between the patient subgroups identified for fingolimod and natalizumab and although an analysis is undertaken using the wider population, the results of these analysis may not reflect the true cost-effectiveness of these treatments and dimethyl fumarate in the two sub-populations. The sub-group analysis was not possible due to a lack of dimethyl fumarate data. The natural history data were based on the dimethyl fumarate trial population where possible. The same considerations regarding the population apply, and this seems largely appropriate given the alternatives of using the London Ontario or MS survey datasets. Little has been published about the London Ontario population and the MS survey population includes SPMS and PPMS patients, with only 35.5% of the MS survey population representing the relevant RRMS population.

There is greater concern about the utility and cost estimates. The utility decrements for SPMS compared to RRMS and for relapse compared to no relapse were based on the MS survey population, which as discussed above is not exactly the population of the scope. The cost estimates for the EDSS states and relapse were also based on the MS survey population. Furthermore, different analyses

based on the same dataset have produced significantly different results. The ERG explores these issues in Section 6. A few of the resource use estimates did not correspond with the experience of the ERG clinical advisor and this is also explored in section 6.

There is uncertainty around the long-term treatment effect of DMTs for MS and associated side-effects. The cost-effectiveness of dimethyl fumarate compared to a cost-effecitveness threshold of £30,000 per QALY did not change given the alternative waning effect assumptions tested by the manufacturer in sensitivity analyses.

In general, the cost-effectiveness outcomes reflect the increased effectiveness of dimethyl fumarate in terms of 3 months sustained disease progression up the EDSS scale and of annualised relapse rates when compared to the comparators except for natalizumab; and the high list price for dimethyl fumarate compared to the interferon drugs and glatiramer acetate. Using the list price for dimethyl fumarate, dimethyl fumarate is not cost-effective. When the manufacturer's proposed PAS price is used for dimethyl fumarate compared to the list prices for other drugs it becomes cost-effective, but this analysis is limited due to the inclusion of the discounted price for only one drug when discounted prices are available to the DoH for all comparators.

6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.1 Overview

This section presents additional ERG analyses exploring alternative model assumptions. For the analyses in this section, there is a technical appendix, section 10.2, which gives details to changes to the manufacturer's model. Although probabilistic sensitivity analysis (PSA) cost-effectiveness results are calculated for different price assumptions in this section, for results for the remaining sensitivity and scenario analyses are presented using deterministic results as the PSA results took too long to compute.

The manufacturer conducted a number of one-way and two-way sensitivity analyses and scenario analyses, which have been discussed in section 5. As outlined, the ERG considered that there were limitations to these sensitivity analyses. This section details the ERG further exploration of these and a number of alternative data estimates which were identified by the ERG. These analyses will be used to investigate the robustness of the results presented. In addition, ERG corrections and adjustments to the manufacturer's base case model are discussed. Changes to the cost-effectiveness results are explored for alternative assumptions made for the following scenarios:

- Treatment costs: drug prices;
- Treatment costs: monitoring costs;
- Relative importance of treatment outcomes;
- Progression: baseline population distribution;
- Progression: treatment discontinuation;
- Progression: transition rates to SPMS;
- Progression: EDSS state utilities;
- Progression: EDSS state costs;
- Relapse: placebo relapse rates;
- Relapse: relapse costs;
- Adverse events: incidence;
- Adverse events: different influenza utilities.

The ERG identified one error in the model. This was the utility estimate for serious and non-serious flu-like symptoms. In the manufacturer's model, the utility estimate was 0.3129 for both serious and non-serious flu-like symptoms. In contrast, the manufacturer's submission (Table 61, P200, section 7.4.8) stated that the utilities should be the same as those for a non-serious influenza event, which was 0.63. The ERG did not correct this in the analyses for this section as the ERG considered 0.3129 to be an appropriate estimate. This is discussed in section 5.2.9.

6.2 Additional ERG analyses

6.2.1 Treatment costs

6.2.1.1 Drug prices

The ERG had concerns regarding the costing scenarios presented. The base case analysis in the manufacturer's submission included the list prices of all the drugs. The manufacturer then presented results of an analysis that included the manufacturer's proposed PAS price and the list price for all other drugs. The ERG considers it more appropriate to compare a discounted price for one drug with the discounted price for another drug, where a discounted price is available.

Risk sharing scheme prices for Rebif 22µg, Rebif 44µg, Avonex, Betaferon and glatiramer acetate were published in a Circular by the Department of Health in 2002.[13] These were also used in a sensitivity analysis within the fingolimod submission. The natalizumab submission did not present a proposed PAS different to the list price of £14,690,[6] so the ERG assumes there is no PAS price for natalizumab. The fingolimod submission did propose a PAS price.[7] This price is not publicly available.

The ERG conducted an analysis adjusting the manufacturer's base case model to include the proposed PAS price for dimethyl fumarate ([redacted]); the risk sharing scheme prices for Rebif 22µg, Rebif 44µg, Avonex, Betaferon and glatiramer acetate; and 35% and 53% reductions for fingolimod. [redacted] The prices used are presented in Table 42.

The pairwise ICERs of dimethyl fumarate compared to each comparator using the discounted prices are compared to the results using the list prices in Table 42. Using the discounted prices, dimethyl fumarate is more cost-effective compared to each comparator than using the list prices. A full incremental cost-effectiveness analysis was then conducted where the drugs are ordered according to increasing cost; the dominated and extendedly dominated drugs are excluded from the calculations; and the ICERs are calculated for the remaining drugs compared to the next most costly drug. The results from a full incremental cost-effectiveness analysis on the deterministic cost and QALY results show that glatiramer acetate is the next most cost-effective drug and that the deterministic ICER for dimethyl fumarate is therefore £36,511 per QALY.

Table 42: The list prices compared to the alternative discounted prices for each drug

	List prices (£)	Discounted prices (£)
Dimethyl fumarate	17,900	[redacted]
Rebif 22µg	8,149	7,513
Rebif 44µg	10,608	8,942
Avonex	8,531	8,502
Glatiramer acetate	6,841	5,823
Fingolimod (35% reduction)	19,176	12,464
Fingolimod (53% reduction)	19,176	9,109
Natalizumab†	14,690	14,690
Betaferon	7,265	7,259

†: There is no discounted price for natalizumab

Table 43: The deterministic pairwise cost-effectiveness results using the list prices for all drugs and discounted prices (where possible) for all drugs

	ICER of DF versus comparator	
	List prices	Discounted prices
Rebif 22 µg	142,283	26,026
Rebif 44 µg	122,105	7,289
Avonex	136,452	DF dominates
Glatiramer acetate	159,295	36,511
Fingolimod (35% red)	DF dominates	DF dominates
Fingolimod (53% red)	DF dominates	DF dominates
Natalizumab†	(173,745)‡	(534,04)
Betaferon	106,127	DF dominates

†: There is no discounted price for natalizumab; ‡: brackets indicate the ICER reflects the reverse comparison, i.e. natalizumab versus dimethyl fumarate

As the model is non-linear, the mean ICER calculated using probabilistic sensitivity analysis is the most appropriate outcome to present. However, it should be noted that probability distributions were not assigned to a number of parameters, as discussed in section 5.2.12.5. As a result, whilst the probabilistic results are the most meaningful the full impact of the uncertainty has not been appropriately characterised in the results of the analysis. The manufacturer claimed in the response to the points of clarification that, when running the adapted rate ratio model with additional parameter distributions, the results were not stable. The ERG tested the stability of the PSA results for the model in the manufacturer's original submission. Using 10,000 iterations, the probabilistic sensitivity results seem stable. Eight analyses were run for dimethyl fumarate compared to glatiramer acetate. The range of results was £49,332 to £50,855 with a mean result of £50,051. The range of results is relatively

insignificant when compared to the difference in deterministic and probabilistic results as shown in Table 45 below.

First, a full incremental cost-effectiveness analysis using the list prices for all drugs is conducted and presented in Table 44 using PSA costs and QALYs.

Secondly, the deterministic and probabilistic pairwise ICER results are compared for each comparator using discounted prices for all drugs where possible in Table 45.

Thirdly, the PSA full incremental cost-effectiveness results using discounted prices where possible for all drugs are presented in Table 46.

The ICER for dimethyl fumarate is higher compared to each comparator when using the PSA results than when using the deterministic results, but represents a less biased approximation of the ICER. When conducting a full incremental cost-effectiveness analysis, glatiramer acetate remains the relevant comparator for dimethyl fumarate, as it is the next best alternative, so the ICER for dimethyl fumarate is £49,687.

Table 44: The probabilistic cost-effectiveness results using the list prices for all drugs

	Costs (£)	QALYs	ICER
Rebif 22	234,103	5.47	-
Glatiramer acetate	234,449	5.50	11,197
Avonex	239,543	5.49	Dominated
Betaferon	239,919	5.44	Dominated
Rebif 44	242,289	5.49	Dominated
Dimethyl fumarate	270,230	5.68	200,117
Fingolimod	281,251	5.51	Dominated
Natalizumab	285,353	5.75	214,815

The ERG proposes that the analysis based on the discounted prices better reflects the cost-effectiveness of these treatments for the NHS. The ERG therefore presents all additional analysis based on these discounted prices.

Table 45: The pairwise deterministic compared to probabilistic cost-effectiveness results using discounted prices (where possible) for each drug

	ICER of DF versus comparator	
	Deterministic (Discounted prices)	PSA (Discounted prices)
Rebif 22	26,026	34,065
Rebif 44	7,289	11,963
Avonex	DF dominates	114
Glatiramer acetate	36,511	49,687
Fingolimod (35% red)	DF dominates	DF dominates
Fingolimod (53% red)	DF dominates	DF dominates
Natalizumab	(534,047)	(691,373)
Betaferon	DF dominates	DF dominates

Table 46: The probabilistic full incremental cost-effectiveness results using discounted prices (where possible) for each drug

	Cost (£)	QALY	ICER
Glatiramer acetate	[redacted]	[redacted]	
Rebif 22	[redacted]	[redacted]	Dominated
Rebif 44	[redacted]	[redacted]	Dominated
Avonex	[redacted]	[redacted]	Dominated
Dimethyl fumarate	[redacted]	[redacted]	49,738
Betaferon	[redacted]	[redacted]	Dominated
Fingolimod (53%)	[redacted]	[redacted]	Dominated
Natalizumab	[redacted]	[redacted]	(407,367)

6.2.1.2 Monitoring costs

The annual monitoring costs of treatment are substantial treatment-related costs. As discussed in section 5.2.10.1, the annual costs after the first year of treatment range from £700 to £1300 for different drugs as presented in Table 64 in section10.2.2. This varies according to the number of neurology visits, MS nurse visits and MRI scans required. The ERG produced an alternative set of resource assumptions listed in Table 36 in section 5.2.10.1. The details are also presented in the Appendix, section 10.2.2. The deterministic pairwise ICER results from making these alternative assumptions and including the reduced prices for all drugs are reported in Table 47.

In addition to some exploration around the impact of monitoring resource use, the ERG also investigated the appropriateness of the unit costs associated with these resources. The cost of a

neurology visit was assumed to be a day case admission in the manufacturer's submission. It is not clear if that is appropriate so an alternative assumption of the cost of a visit to a neurology specialist was added to the other ERG monitoring resource assumptions (Appendix, section 10.2.2). The deterministic pairwise ICER results including the neurology visit cost are also presented in Table 47. The ICERs from the analysis incorporating all the alternative ERG assumptions are not very different to those from the analysis with the original manufacturer's assumptions.

Table 47: The pairwise deterministic cost-effectiveness results for each dimethyl fumarate compared to each comparator using discounted prices for all drugs (where possible) in the ERG base case, alternative ERG monitoring assumptions, and different monitoring assumptions plus a different cost of neurology

	Base case (£/QALY)	ERG monitoring assumptions (£/QALY)	ERG monitoring assumptions and reduced cost of neurology visit (£205)- (£/QALY)
Rebif 22 µg	26,026	34,893	28,168
Rebif 44 µg	7,289	17,091	9,895
Avonex	DF dominates	7,084	491
Glatiramer acetate	36,511	43,874	37,791
Fingolimod (35%)	DF dominates	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates	DF dominates
Natalizumab	(534,047)	(526,405)	(524,256)
Betaferon	DF dominates	2,624	DF dominates

The ERG's clinical expert highlighted the need for MS nurse visits for injectable treatments. The analysis above assumes the nurse visits are required by all DMTs; this may not be the case for the treatments where the injection is not self-administered (natalizumab) or where an injection is not necessary (fingolimod and natalizumab). Therefore, the ERG present two additional scenarios, one where all injectable treatments incur nurse visits (Rebif 22, Rebif 44, Avonex, glatiramer acetate and natalizumab) and one where only self-injectable treatments incur nurse visits (Rebif 22, Rebif 44, Avonex, glatiramer acetate and natalizumab). The results of the analyses for all injectables are presented in Table 48. Natalizumab was the only drug that was not self-injectable. The very high ICER for natalizumab versus dimethyl fumarate hardly changed when alternative nurse visit assumptions were made.

Table 48: The pairwise deterministc ICERs using discounted prices for all drugs (where possible) where MS nurse visit is included for all injectables

	Base case neurology cost (£/QALY)	Reduced neurology cost (£/QALY)
Avonex	4,733	DF dominates
Rebif 22 µg	32,810	26,085
Rebif 44 µg	14,771	7,576
Betaferon	753	DF dominates
Fingolimod (35%)	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates
Glatiramer acetate	41,513	37,477
Natalizumab	(532,472)	(530,323)

6.2.2 Relative importance of treatment outcomes

The manufacturer conducted many one-way and two-way sensitivity analyses, and consistently it was variation in the risk of disability progression that had the greatest impact on the ICER. In order to get an overview of the relative importance of the different outcomes to the results, the ERG conducted the following scenario analyses:

- The relative annualised relapse rate for dimethyl fumarate and the comparator compared to placebo was set to 1;
- The relative risk of progression of dimethyl fumarate and the comparator compared to placebo was set to 1;
- No waning effect after 2 years was assumed;
- Complete waning effect after 2 years was assumed; and
- No adverse events was assumed

The deterministic pairwise ICER results from making these alternative assumptions and including the reduced prices for all drugs are reported in Table 49. The utilities and costs of adverse events are incorporated in the model independently of the discontinuation of treatment. Assuming that there are no dis-utilities or costs associated with adverse events has relatively little effect on the ICER. Assuming equal relapse rates has more of an effect on the ICER but the impact is still relatively small. The effect of treatment on disability progression and treatment waning have by far the greatest effects on the results. Although the robustness of the results to alternative assumptions for parameters related to relapse rates and adverse events is explored below, given the uncertainty in many of the estimates the focus of the following analyses on parameters related to disease progression as that is clearly the driver of the model.

Table 49: The deterministic pairwise cost-effectiveness results using discounted prices for all drugs (where possible) for alternative treatment outcome assumptions

		DF versus the comparator		Waning effect after 2 years		No adverse events
	Base case	ARR† = 1	RRP‡ =1	None	Complete	
Rebif 22µg	26,026	34,347	285,965	14,850	128,874	30,563
Glatiramer acetate	36,511	40,998	818,131	25,502	139,390	37,818
Avonex	Dominates	6,534	26,288	Dominates	28,862	Dominates
Betaferon	Dominates	Dominates	40,749	Dominates	12,296	Dominates
Rebif 44µg	7,289	13,823	118,014	1,066	61,187	10,470
Fingolimod (35%)	Dominates	Dominates	Dominates	Dominates	Dominates	Dominates
Fingolimod (53%)	Dominates	Dominates	Dominates	Dominates	Dominates	Dominates
Nataluzimab	(534,047)	(609,157)	Dominates	(360,812)	(1,451,485)	(468,735)

†: ARR: Annualised relapse rate; ‡: RRP: Relative risk of progression

6.2.3 Disease progression

6.2.3.1 Baseline population distribution

The model structure involves patients transitioning between RRMS and SPMS and between EDSS states within each of those. The higher EDSS states incur progressively greater disutility and cost. It is not clear what distribution of the RRMS population across EDSS states corresponds with the scope population. The baseline distribution is based on the dimethyl fumarate trials, which has a slightly, but not significantly different population to that in the scope as discussed in section 5.2.2. The importance of the distribution across EDSS states to the results is tested simply by making different assumptions based on 100% of the population starting in different EDSS states. The results are presented in Table 50.

In general, the lower the EDSS state in which a patient starts, the less cost-effective dimethyl fumarate compared to the comparator. The reason is that if the population is grouped in lower EDSS states at the start of the model, there is less immediate benefit from a reduction in disability progression due to drugs. Patients incur treatment costs, but the costs and disutilities associated with the lower EDSS states are less. In Table 50 the reduction in the ICER compared to glatiramer acetate reduced by £6,986 from EDSS 4 to 5. Excluding withdrawal from the model has little effect on this as withdrawal rates are similar. However, assuming no waning effects the reduction is only £3,531; and assuming no discounting the reduction is only £3,432. This is because more patients develop higher EDSS states quicker and future benefits are valued more.

The distribution of the population across the first 5 EDSS states has little impact on the ICER and the ERG base case deterministic cost-effectiveness results are not very different to assuming that the whole population has the baseline characteristic of an EDSS state of zero. Therefore, the slight

difference in the trial population compared to the general RRMS population does not have a significant effect on the cost-effectiveness conclusions in terms of EDSS states.

Table 50: The deterministic pairwise ICERs for dimethyl fumarate compared to each comparator using discounted prices, assuming 100% of the population starts the model in different EDSS states 0 to 5

ICER of Dimethyl fumarate v comparator (discounted prices)	Original population distribution	100% of population in EDSS state					
		0	1	2	3	4	5
Rebif 22µg	26,026	**28,389**	**27,218**	**27,019**	**25,829**	**24,882**	**17,469**
Glatiramer acetate	36,511	**36,606**	**35,420**	**36,356**	**36,722**	**38,967**	**31,981**
Avonex	Dominates	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**
Betaferon	Dominates	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**
Rebif 44µg	7,289	**10,037**	**9,088**	**8,396**	**7,073**	**5,066**	**Dominates**
Fingolimod (35%)	Dominates	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**
Fingolimod (53%)	Dominates	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**	**Dominates**
Nataluzimab	(534,047)	**(523,695)**	**(514,982)**	**(531,611)**	**(534,096)**	**(567,858)**	**(508,562)**

6.2.3.2 Treatment discontinuation

In the manufacturer's model, patients may come off treatment for three reasons:

- Discontinuation due to adverse events;
- Progressing to EDSS state 7; and
- Developing SPMS.

The effect on the ICER of making the following assumptions is evaluated:

- The discontinuation risk reduced by 50% after 2 years of treatment for both dimethyl fumarate and the comparator;
- That there is no discontinuation risk after 2 years of treatment for both dimethyl fumarate and the comparator; and
- The 95% lower or upper limits of the confidence interval for the relative risk of discontinuation of treatment due to any cause.

The deterministic pairwise ICER results from making these alternative assumptions are reported in Table 51. Changing the discontinuation risk for both comparators at the same time after 2 years of

treatment has little effect on the ICER as you would expect. When the relative discontinuation risk is changed there is a much greater effect on the ICER, but it does not change the ERG base case results significantly.

After discontinuation the model assumes that patients receive no treatment. This means that in any treatment arm of the model, a proportion of the patients are on treatment and a proportion are on no treatment – which in clinical practice the ERG believes would be 'best supportive care' or an alternative DMT. Due to discontinuation the proportion of patients off-treatment increases over time. The results clearly indicate that had 'best supportive care' been included formally as a comparator, none of the DMTs being assessed would be considered cost-effective compared to NICE's conventional threshold for cost-effectiveness. When a patient discontinues treatment, they receive placebo or 'best supportive care' and make a quicker progression through EDSS states. The switch from treatment to no treatment leads to a significant reduction in treatment costs combined with a small reduction in QALYs. This creates a situation where the ICER behaves in an unintuitive manner; that is, the more patients that discontinue treatment, the more cost-effective the drug appears. If 'best supportive care' were one of the comparators included in the scope, then the ICER of a drug would never be less than the ICER of the drug compared to 'best supportive care'.

In Table 52, the effect of increasing the discontinuation rates for both diemthyl fumarate and Rebif 22 µg on the total costs and QALYs for each intervention and the ICER of dimethyl fumarate compared to Rebif 22 µg is presented. As the treatment costs for dumethyl fumarate are higher than those for Rebif 22 µg, the total costs reduce more as patients come off treatment and the reduction in QALYs does not offset that.

Glatiramer acetate remains the relevant comparator (i.e. the next best alternative) in a full incremental cost-effectiveness analysis in all scenarios.

As mentioned in section 5, it seems more likely that those patients who stop treatment for adverse effects will move to another active treatment if an alternative is available with a difference adverse event profile. Further, it is also possible that those patients who progress to SPMS will receive some form of active treatment which has not been included in the model due to limitations in data and scope.

Table 51: The deterministic pairwise ICERs of dimethyl fumarate versus each comparator using the upper limit (UL) or lower limit (LL) of the confidence interval of the relative risk of discontinuation of the comparator versus dimethyl fumarate

		ICER of DF versus comparator			
	Base case	50% discontinuation rate after 2 years	0% discontinuation rate after 2 years	95% LL of CI for relative discontinuation risks[Ⅱ]	95% UL of CI for relative discontinuation risks
Rebif 22 µg	26,026	27,988	27,594	37,980	DF dominates
Rebif 44 µg	7,289	7,775	4,770	22,054	DF dominates
Avonex	DF dominates	DF dominates	1,430	16,441	DF dominates
Glairamer acetate	36,511	40,633	48,436	40,546	31,367
Fingolimod (35%)	DF dominates	DF dominates	DF dominates	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates	DF dominates	DF dominates	DF dominates
Natalizumab	(534,047)	(590,359)	(700,231)	DF dominates	(334,095)
Betaferon	DF dominates	327	10,240	8,562	DF dominates

Ⅱ: Dimethyl fumarate relative to the comparator, so at the lower limit of the confidence interval Dimethyl fumarate has a lower discontinuation risk relative to the comparator.

Table 52: The change in total costs and QALYs for dimethyl fumarate and Rebif 22 µg and the ICER of dimethyl fumarate compared to Rebif 22 µg when discontinuation rates are assumed to be the same and are increased in 5% increments

	Dimethyl fumarate		Rebif 22 µg		ICER (dimethyl fumarate vs Rebif 22 µg)
Discontinuation rates	Total cost	QALY	Total cost	QALY	
5%	251,083,905	5,906	246,695,004	5,624	15,587
10%	242,009,291	5,777	239,102,769	5,552	12,925
15%	236,227,348	5,682	234,162,445	5,498	11,192

6.2.3.3 Change in transition rates to SPMS

The transition rate to SPMS is independent of treatment, but as no treatment is received once this transition takes place patients, in effect, are discontinuing treatments. The transition rates to SPMS were derived from the London Ontario dataset. As it is not clear how appropriate the London Ontario population is for the scope population, the ERG conducted sensitivity analysis around the transition rate to SPMS for each EDSS state. The transition rates were increased and decreased by 50% in an attempt to assess the robustness of the results in relation to these data.

The deterministic pairwise ICER results from making these alternative assumptions and including the reduced prices for all drugs are reported in Table 53. The increase in patients on placebo in both

treatment arms results in a small reduction in the ICER for dimethyl fumarate compared to the comparator. The explanation for this was discussed in section 6.2.3.2.

Table 53: The deterministic pairwise cost-effectiveness results for dimethyl fumarate compared to each comparator using discounted prices for all drugs and assuming a 50% increase of decrease in the transition rate to SPMS from RRMS in every EDSS state

	ICER of dimethyl fumarate versus comparator		
	Base case	Transition rate for each EDSS state (0-8) to SPMS increased by 50%	Transition rate for each EDSS state to SPMS reduced by 50%
Rebif 22	26,026	22,356	30,201
Rebif 44	7,289	4,421	10,356
Avonex	DF dominates	DF dominates	DF dominates
Glatiramer acetate	36,511	34,345	39,568
Fingolimod (35%)	DF dominates	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates	DF dominates
Natalizumab	(534,047)	(502,312)	(576,594)
Betaferon	DF dominates	DF dominates	DF dominates

6.2.3.4 Utilities for EDSS states

Disease progression has a significant effect on the ICER because of the increasing disutility and cost of higher RRMS EDSS states and SPMS states relative to RRMS states. The absolute values of the utilities for the different EDSS states were derived from the dimethyl fumarate trials and were assumed in the manufacturer's model to apply to RRMS patients who are not in a relapse state. This may underestimate the utilities as some patients will have been in a relapse state. Consequently, in an attempt to explore this uncertainty the ERG substituted the utilities for the EDSS states based on the natalizumab submission.[6] This utility dataset has the same utility decrements associated with relapse versus no relapse and SPMS versus RRMS, but the absolute level of utility is slightly higher than in the manufacturer's submission.

As discussed in section 4.2.2, the dimethyl fumarate trial population is not exactly the same as the scope population. As a result, the absolute utility values from the population of the MS survey, as reported in Orme et al.[39] were also used to explore the effect on the results. This is also the population that provided the utility decrements for relapse versus no relapse and SPMS versus RRMS. This is not a preferred population, but it tests the sensitivity of the results to different baseline utility values. The MS survey population is discussed in section 5.2.

The deterministic pairwise ICER results from making these alternative assumptions are reported in Table 54. The different utility estimates do not have a significant impact on the results. Glatiramer acetate remains the relevant comparator (i.e. the next best alternative) in a full incremental cost-effectiveness analysis.

Table 54: The deterministic pairwise cost-effectiveness results of dimethyl fumarate compared to each comparator using alternative EDSS utility estimates from Orme et al and TA127[6, 39]

	ICER of dimethyl fumarate versus comparator		
	Base case	Orme	MS survey (TA127)
Rebif 22 µg	26,026	22,271	26,952
Rebif 44 µg	7,289	6,404	7,540
Avonex	DF dominates	DF dominates	DF dominates
Glairamer acetate	36,511	34,427	37,952
Fingolimod (35%)	DF dominates	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates	DF dominates
Natalizumab	-534,047	-344,180	-555,052
Betaferon	DF dominates	DF dominates	DF dominates

6.2.3.5 Alternative EDSS State costs

The cost of different EDSS states will also affect the impact of disease progression on the ICER. The costs associated with different EDSS states were derived from the MS survey. As discussed in section 5.2.10.1, the ERG has identified three different cost estimates based on the same MS survey[6, 7] and it is not completely clear why the results are different, although it is known that the unit costs vary slightly between two of the analyses and Tyas et al distinguished between medical and non-medical costs.[45] The ERG therefore inflated the different cost estimates to the year 2012 and evaluated the effect of these different costs on the ICER.

The deterministic pairwise ICER results from making these alternative assumptions and including the reduced prices for all drugs are reported in Table 55. The higher EDSS cost estimates from Tyas et al decrease the ICER of dimethyl fumarate because it increases the benefit from a reduction in disability progression. The lower cost estimates from the natalizumab submission (TA127) has the opposite effect. The different EDSS state cost estimates do not have a significant impact on the results. Glatiramer acetate remains the relevant comparator in a full cost-effectiveness analysis.

Table 55: The deterministic pairwise ICERs for dimethyl fumarate compared to each comparator for the ERG base case with discounted prices for all drugs, and different EDSS state cost estimates from TA127 and Tyas et al[6, 46]

	Base case (£/QALY)	TA127 EDSS State costs (£/QALY)	Tyas et al EDSS states costs (medical and non-medical)- (£/QALY)
Rebif 22 µg	26,026	28,575	21,981
Rebif 44 µg	7,289	9,763	3,354
Avonex	DF dominates	DF dominates	DF dominates
Glatiramer acetate	36,511	39,248	32,157
Fingolimod (35%)	DF dominates	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates	DF dominates
Natalizumab	(534,047)	(537,065)	(529,162)
Betaferon	DF dominates	DF dominates	DF dominates

6.2.4 Relapse

It was indicated in section 6.2.2 that relapse rates were not expected to have as a great an impact on the model results as disease progression. Nevertheless, uncertainty in the placebo relapse rate and in the cost of the relapse motivated further analyses by the ERG.

6.2.4.1 Placebo relapse rates

The relapse rates for patients on placebo were based on the relapse rates for patients in the 12 months prior to randomisation in the dimethyl fumarate trials. Since a relapse in the last 12 months is not population criterion in the scope it is possible that the relapse rates from the trial are too high for the general RRMS population. Although the MS survey is not considered the same population as the scope either (see section 5.2), the relapse rates are closer to the ERG clinical advisor's estimate of 0.8 per year. The sensitivity of the results to these alternative values was also tested.

The deterministic pairwise ICER results from making these alternative assumptions and including the reduced prices for all drugs are reported in Table 56. The different placebo relapse rate estimates do not have a significant impact on the results. Glatiramer acetate remains the relevant comparator (i.e. the next best alternative) in a full cost-effectiveness analysis.

Table 56: The deterministic pairwise ICERs for dimethyl fumarate compared to each comparator using the ERG base case of discounted prices for all drugs and alternative placebo relapse rates from the MS survey

	ICER of dimethyl fumarate versus comparator	
	Base case	MS survey natural history relapse rates RRMS and SPMS
Rebif 22 µg	26,026	29,698
Rebif 44 µg	7,289	10,151
Avonex	DF dominates	1,462
Glairamer acetate	36,511	38,356
Fingolimod (35% red)	DF dominates	DF dominates
Fingolimod (53% red)	DF dominates	DF dominates
Natalizumab*	(534,047)	(567,299)
Betaferon	DF dominates	DF dominates

Alternative Relapse costs

As previously discussed, the ERG identified four different cost estimates for a relapse. Three of these were derived from the same MS survey source.[6, 7, 46] The costs range from £208 to £3,039 in their respective price years. Furthermore, the ERG clinical advisor estimated that only 20% of RRMS patients with a relapse would be admitted to hospital making the estimate of £2,028 seem too high. It is assumed here that the 80% would incur no cost although on average some cost would be incurred. The ERG conducted four different analyses based on these considerably different cost estimates. The details are presented in the Appendix 10.2.4.

The deterministic pairwise ICER results from making these alternative assumptions are reported in Table 57. The different relapse cost estimates do not have a significant impact on the results. Glatiramer acetate remains the relevant comparator in a full cost-effectiveness analysis.

Table 57: The deterministic pairwise ICERs for dimethyl fumarate compared to each comparator using different relapse cost estimates as detailed in Appendix 10.2.4.2

	Base case	Relapse cost used in the Natalizumab submission[6]	Relapse cost used in the Fingolimod submission (cost of an inpatient admission)[7]	Tyas et al relapse costs[46]	20% inpatient admission, 80% not presenting for treatment
	£2,028/relapse	£280.41/relapse	£3,039/relapse	£1,996.09/relapse	£607.8/relapse
Rebif 22µg	26,026	31,446	22,892	26,126	30,431
Rebif 44µg	7,289	11,824	4,667	7,372	10,974
Avonex	DF dominates	4,372	DF dominates	DF dominates	3,053
Glatiramer acetate	36,511	38,923	35,116	36,555	38,471
Betaferon	DF dominates	DF dominates	DF dominates	DF dominates	DF dominates
Fingolimod (35%)	DF dominates	DF dominates	DF dominates	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates	DF dominates	DF dominates	DF dominates
Natalizumab	(534,047)	(550,365)	(524,610)	(534,347)	(547,309)
Betaferon	26,026	DF dominates	DF dominates	DF dominates	DF dominates

6.2.5 Adverse events

It was indicated in section 6.2.2 that the costs and utilities of adverse events were not expected to have as great an impact on the model results as disease progression. Nevertheless, uncertainty in the adverse event incidence rate and in the utility of influenza motivated further analyses.

6.2.5.1 The incidence of adverse events

It should be noted that one of the criterion for an adverse event to be included in the analysis was that it was an adverse event that occurred in a dimethyl fumarate trial. This assumption is conservative with respect to dimethyl fumarate as there are likely other adverse events related to the other drugs not included in this analysis. However, as discussed in section 5.2.7 there are adverse effects associated with dimethyl fumarate which did not occur within the trial.

Although the manufacturer conducted an MTC of the included adverse events across all of the comparators, the relative risk results were not used in the economic evaluation. Instead, incidence rates across the trials for each comparator were used. No explanation was given for this. The ERG therefore took the results data and calculated new incidence rates for each treatment for each adverse event as detailed in the Appendix, section 10.2.5.

The deterministic pairwise ICER results from making these alternative assumptions and including the reduced prices for all drugs are reported in Table 58, alongside the assumption of no adverse events.

Table 58: The deterministic pairwise ICERs for dimethyl fumarate compared to each comparator assuming that the adverse event rates were obtained from the MTC, that there were no adverse events, and compared to the base case where event rates were calculated independently

	Base case	Adverse events derived from MTC	No adverse events
Rebif 22µg	26,026	32,819	30,563
Glatiramer acetate	36,511	37,176	37,818
Avonex	Dominates	Dominates	Dominates
Betaferon	Dominates	Dominates	Dominates
Rebif 44µg	7,289	10,884	10,470
Fingolimod (35%)	Dominates	Dominates	Dominates
Fingolimod (53%)	Dominates	Dominates	Dominates
Nataluzimab	(534,047)	(471,763)	(468,735)

1: ARR: Annualised relapse rate; 2: RRP: Relative risk of progression; the negative ICER indicates that Dimethyl fumarate is less effective and cheaper than the comparator

6.2.5.2 Different influenza utilities

The dis-utilities estimated for influenza and flu-like symptoms appeared high. These dis-utilities were derived from the Van Hoek et al (2011) study in which EQ-5D results were presented for baseline(confirmed and unconfirmed flu) and worst day (confirmed and unconfirmed flu).[49] In addition, overall QALY losses over the duration of influenza were also reported. Van Hoek utilities are presented in Table 59.

Table 59: Van Hoek utility results

Van Hoek results	Confirmed flu	Unconfirmed flu
EQ-5D baseline	0.96	0.97
EQ-5D worst day	0.29	0.34
Overall QALY loss	0.008	0.0075

The manufacture undertook the following calculations to derive utility estimates.

Decrement for non-serious influenza, non-serious and serious flu like symptoms = EQ-5D unconfirmed baseline (0.97) minus EQ-5D unconfirmed worst day (0.34) = decrement (0.63)

Decrement for serious influenza = EQ-5D confirmed baseline (0.96) minus EQ-5D confirmed worst day (0.29) = decrement (0.67)

Given that overall QALY lost is also reported the ERG felt that using that value to obtain a utility decrement would be plausible.

Decrement for non-serious influenza, non-serious and serious flu like symptoms = overall QALY loss unconfirmed flu (0.0075) divided by the duration of influenza (8.75 days) multiplied by days in a year (365) = 0.334.

Decrement for serious influenza = overall QALY loss confirmed flu (0.008) divided by the duration of influenza (8.75 days) multiplied by days in a year (365) = 0.313.

The impact of these alternative disutility estimates are presented in table 60.

Table 60: Alternative utility estimates for flu-like symptoms and influenza

	ICER of DF versus comparator	
	Base case	0.313 utility for flu-like symptoms and non-serious influ; 0.334 for serious influenza
Rebif 22 µg	26,026	25,919
Rebif 44 µg	7,289	7,275
Avonex	DF dominates	DF dominates
Glairamer acetate	36,511	36,504
Fingolimod (35%)	DF dominates	DF dominates
Fingolimod (53%)	DF dominates	DF dominates
Natalizumab	(534,047)	(524,352)
Betaferon	DF dominates	DF dominates

6.3 Conclusions from ERG analyses

Overall the modelling approach adopted by the manufacturer was appropriate. However, some uncertainties remain around the data used to inform the progression of disease. The single biggest factor that affects the absolute incremental cost-effectiveness results is the prices of the drugs and the comparators included. The ERG considers the appropriate analyses to be where the list prices are used for every drug and where the reduced prices (known or estimated through sensitivity analyses) are used for every drug. The ERG also considers the ICER calculated from PSA results to be preferable to the deterministic ICER. Although it would have been preferable to have had distributions assigned to all parameters, rather than some.

After discontinuation the model assumes that patients receive no treatment. This means that in any treatment arm of the model, a proportion of the patients are on treatment and a proportion are on no treatment – which in clinical practice the ERG believes would be 'best supportive care' or an alternative DMT. Due to discontinuation the proportion of patients off-treatment increases over time. When a patient discontinues treatment for any reason, the model assumes they receive no treatment

and therefore make a quicker progression through EDSS states. The switch from treatment to no treatment leads to a significant reduction in treatment costs combined with a small reduction in QALYs. This creates a situation where the ICER behaves in an unintuitive manner; that is, the more patients that discontinue treatment, the more cost-effective the drug appears. This situation is unlikely to reflect clinical practice.

Parameters related to disease progression had the greatest influence on the cost-effectiveness results. The main parameter of the model, the relative risk of 3 months sustained progression, was however considered to have been derived from an adequate MTC analysis, and the ERG conducted no further analyses on this parameter. There is uncertainty around the relative treatment effects, and the manufacturer's own sensitivity analyses revealed that plausible changes in the relative effects within the confidence intervals changed the cost-effectiveness conclusion for dimethyl fumarate.

There is also considerable uncertainty around the relative discontinuation risks, and these have a significant impact on the ICER estimates, although dimethyl fumarate is never cost-effective compared to glatiramer acetate when using the PSA results across the range of relative discontinuation risks.

Although there is also considerable uncertainty in many other parameter estimates, no alternative estimates were identified by the ERG that had a significant impact on the results. When the list prices are used the ICER for Dimethyl fumarate is at least £159,295. When the reduced prices are used, the ICER for dimethyl fumarate is £49,687.

7 End of life

Dimethyl fumarate does not meet the end of life criteria.

8 Overall conclusions

Evidence from two good quality RCTs demonstrates that dimethyl fumarate is effective in reducing the relapse rate in RRMS patients compared to placebo over a two year period, though there was some uncertainty regarding the benefit for disability progression. To obtain relative treatment effects with the comparators outlined in the scope an MTC was undertaken. There was some heterogeneity across the trials included in the MTC and therefore some uncertainty surrounding the results obtained. The general RRMS population was outlined in the scope; in addition two subpopulations for which natalizumab and fingolimod have been licensed and recommended were also suggested as relevant populations. No analyses were conducted on these populations and the efficacy, and subsequent cost-effectiveness, of dimethyl fumarate compared to natalizumab and fingolimod in their licensed indications is therefore uncertain.

No previously published cost-effectiveness results were pertinent to the decision problem. All aspects of the *de novo* economic evaluation were consistent with the NICE reference case with the exception that best supportive care was excluded as a comparator. This, however, was consistent with the scope. The driver of the model was disease progression through EDSS states. Across the sensitivity analyses conducted by the manufacturer and ERG, the results were mainly sensitive to variation in the rate of 3 months sustained progression and the price of treatment. Given the data and assumptions included in the model, none of the treatments were cost-effective compared to placebo, and this had the perverse effect that the higher the drop-out rate, the more cost-effective the treatment. The ERG feel that the most plausible ICER lies somewhere in the region of £49,687.

8.1 Implications for research

There is a need for improved long-term data on the natural history of RRMS to inform future cost-effectiveness analyses. In relation to dimethyl fumarate, longer term data on effectiveness and safety is required.

9 References

1. Office for National Statistics. *Population estimates for England and Wales, mid-2011 (2011 census-based)*. ONS; 2012. [cited 2013 July 17]. Available from: http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-262039.

2. Fox RJ, Bethoux F, Goldman MD, Cohen JA. Multiple sclerosis: advances in understanding, diagnosing, and treating the underlying disease. *Cleve Clin J Med* 2006;73:91-102.

3. Rosati G. The prevalence of multiple sclerosis in the world: an update. *Neurol Sci* 2001;22:117-39.

4. Zajicek JP, Ingram WM, Vickery J, Creanor S, Wright DE, Hobart JC. Patient-orientated longitudinal study of multiple sclerosis in south west England (The South West Impact of Multiple Sclerosis Project, SWIMS) 1: protocol and baseline characteristics of cohort. *BMC Neurol* 2010;10:88.

5. National Institute for Clinical Excellence. *Beta interferon and glatiramer acetate for the treatment of multiple sclerosis. Technology appraisal No. 32*. NICE; 2002. [cited 2013 July 17]. Available from: http://www.nice.org.uk/nicemedia/live/11441/32290/32290.pdf.

6. National Institute for Health and Clinical Excellence. *Natalizumab for the treatment of adults with highly active relapsing–remitting multiple sclerosis. NICE technology appraisal guidance 127*. NICE; 2007. [cited 2013 July 17]. Available from: http://www.nice.org.uk/nicemedia/live/11822/36136/36136.pdf.

7. National Institute for Health and Clinical Excellence. *Fingolimod for the treatment of highly active relapsing–remitting multiple sclerosis. NICE technology appraisal guidance 254*. NICE; 2012. [cited 2013 July 17]. Available from: http://www.nice.org.uk/nicemedia/live/13719/58914/58914.pdf.

8. Association of British Neurologists. *Revised (2009) Association of British Neurologists' guidelines for prescribing in multiple sclerosis* London: ABN 2009. Available from: http://www.theabn.org/abn/userfiles/file/ABN_MS_Guidelines_2009_Final(1).pdf

9. 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.

10. 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.

11. Kappos L, Radue EW, O'Connor P, Polman C, 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.

12. Cohen JA, Barkhof F, Comi G, Hartung HP, Khatri BO, Montalban X, et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N Engl J Med* 2010;362:402-15.

13. *Cost effective provision of disease modifying therapies for people with multiple sclerosis. Health Service Circular 2002/004*. Department of Health; 2002. [cited 2013 July 17]. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4012214.pdf.

14. European Medicines Agency. *Guideline on clinical investigation of medicinal products for the treatment of Multiple Sclerosis [draft]*. London: EMA 2012.

15. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.

16. Fox RJ, Miller DH, Phillips JT, Hutchinson M, Havrdova E, Kita M, et al. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis: trial protocol. *N Engl J Med* 2012;367:1087-97.

17. Gold R, Kappos L, Arnold DL, Bar-Or A, Giovannoni G, Selmaj K, et al. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis: trial protocol. *N Engl J Med* 2012;367:1098-107.

18. Fox RJ, Miller DH, Phillips JT, Hutchinson M, Havrdova E, Kita M, et al. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *N Engl J Med* 2012;367:1087-97.

19. Cadavid D, Wolansky LJ, Skurnick J, Lincoln J, Cheriyan J, Szczepanowski K, et al. Efficacy of treatment of MS with IFN(beta)2-1b or glatiramer acetate by monthly brain MRI in the BECOME study. *Neurology* 2009;72:1976-83.

20. 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.

21. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011;69:292-302.

22. Phillips JT, Fox RJ, Selmaj K, Raghupathi K, Yuan H, Novas M, et al. Safety and tolerability of oral BG-12 (dimethyl fumarate) in relapsing–remitting multiple sclerosis (RRMS): interim results from ENDORSE extension study *Neurology* 2013;80 (Meeting Abstracts 1):P01.162.

23. Kappos L, Gold R, Miller DH, Macmanus DG, Havrdova E, Limmroth V, et al. Efficacy and safety of oral fumarate in patients with relapsing-remitting multiple sclerosis: a multicentre, randomised, double-blind, placebo-controlled phase IIb study. *Lancet* 2008;372:1463-72.

24. 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.

25. Kappos L, Li D, Calabresi PA, O'Connor P, Bar-Or A, Barkhof F, et al. Ocrelizumab in relapsing-remitting multiple sclerosis: a phase 2, randomised, placebo-controlled, multicentre trial. *Lancet* 2011;378:1779-87.

26. Saida T, Kira J, Kikuchi S, Itoyama Y, Hao Q. *Phase 2 study of fingolimod efficacy and safety in Japanese patients*. Americas Committee for Treatment and Research in Multiple Sclerosis; 2012.

27. 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.

28. Hurwitz BJ, Jeffery D, Arnason B, Bigley K, Coyle PK, 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.

29. Lublin FD, Cofield SS, Cutter GR, Conwit R, Narayana PA, Nelson F, et al. Randomized study combining interferon and glatiramer acetate in multiple sclerosis. *Ann Neurol* 2013;73:327-40.

30. Ermis U, Weis J, Schulz JB, Techni RW. PML in a patient treated with fumaric acid. *N Engl J Med* 2013;368:1657-58.

31. Sweetser MT, Dawson KT, Bozic C. Manufacturer's response to case reports of PML. *N Engl J Med* 2013;368:1659-61.

32. van Oosten BW, Killestein J, Barkhof F, Polman CH, Wattjes MP. PML in a patient treated with dimethyl fumarate from a compounding pharmacy. *N Engl J Med* 2013;368:1658-59.

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

34. Department of Health. *Reference costs 2011-12*. London: DH; 2012. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213060/2011-12-reference-costs-publication.pdf

35. Department of Health. *Payment by Results in the NHS: tariff for 2012 to 2013*. London: DH; 2012. Available from: https://www.gov.uk/government/publications/confirmation-of-payment-by-results-pbr-arrangements-for-2012-13

36. Curtis L. *Unit costs of health & social care 2012*. The University of Kent, Personal Social Services Research Unit; 2012. [cited 2013 July 17]. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2012/.

37. 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.

38. Kobelt G, Berg J, Lindgren P, Kerrigan J, Russell N, Nixon R. Costs and quality of life of multiple sclerosis in the United Kingdom. *Eur J Health Econ* 2006;7 Suppl 2:S96-S104.

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

40. Cottrell DA, Kremenchutzky M, Rice GPA, Koopman WJ, Hader W, Baskerville J, et al. The natural history of multiple sclerosis: a geographically based study. 5. The clinical features and natural history of primary progressive multiple sclerosis. *Brain* 1999;122:625–39.

41. Kremenchutzky M, Rice GPA, Baskerville J, Wingerchuk DM, Ebers GC. The natural history of multiple sclerosis: a geographically based study. 9. Observations on the progressive phase of the disease. *Brain* 2006;129:584–94.

42. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;33:1444-52.

43. Excellence NIfC. *Multiple sclerosis: management of multiple sclerosis in primary and secondary care. Clinical guideline 8*. NICE; 2003. [cited 2013 July 17]. Available from: http://www.nice.org.uk/nicemedia/pdf/cg008guidance.pdf.

44. Tyas D, Kerrigan J, Russell N, Nixon R. The distribution of the cost of multiple sclerosis in the UK: how do costs vary by illness severity? *Value Health* 2007;10:386-89.

45. Karampampa K, Gustavsson A, Miltenburger C, Tyas D. Treatment experience, burden and unmet needs (TRIBUNE) in MS study: results from the United Kingdom. 2012;18(6 Suppl 2):41–45.

46. Curtis L. *Unit costs of health & social care 2011*. The University of Kent, Personal Social Services Research Unit; 2011. [cited 2013 July 24]. Available from: http://www.pssru.ac.uk/archive/pdf/uc/uc2011/uc2011.pdf.

47. Kingwell E, van der Kop M, Zhao Y, Shirani A, Rieckmann P, H. T. Mortality and predictors of survival in multiple sclerosis: findings from British Columbia, Canada. *Mult Scler* 2010;16(10 Suppl) S11-S12.

48. van Hoek AJ, Underwood A, Jit M, Miller E, Edmunds W. The impact of pandemic influenza H1N1 on health-related quality of life: a prospective population-based study. *PLoS ONE* 2011;6:e17030.

49. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. *Value Health* 2013;16:686-95.

10 Appendices

10.1 Detailed Critique of Search strategies

10.1.1 Search strategy for identification of clinical effectiveness studies (manufacturer's submission section 6.1).

The submission gave detailed descriptions of the search strategies and met NICE requirements. All of the databases specified by NICE in the specification for manufacturer/sponsor submission of evidence (MEDLINE, MEDLINE In-Process, EMBASE, and CENTRAL) were searched.

Additional searches were carried out on the NICE website, ClinicalTrials.gov, and metaRegister of Controlled Trials (mRCT). Reference lists of previous trials and systematic reviews were searched as well as unpublished data from clinical study reports held by the manufacturer.

The following Web sites were searched for conference abstracts that were published from 2009 to 2012: European Committee for the Treatment and Research in Multiple Sclerosis, (ECTRIMS), Americas Committee for Treatment and Research in Multiple Sclerosis, (ACTRIMS), American Academy of Neurology (AAN), American Neurological Association (ANA), European Federation of Neurological Societies (EFNS).

The searches were undertaken to inform the systematic review of dimethyl fumarate and the MTC and aimed to retrieve RCTs relating to the use of dimethyl fumarate, the various beta interferons, glatiramer acetate, fingolimod, natalizumab, and teriflunomide for relapsing remitting multiple sclerosis (RRMS).

The terms used for each search facet were appropriate. The search strategies were structured using a combination of subject indexing and free text search terms; and search facets were correctly combined using Boolean operators. The choice of MeSH indexing terms was not entirely consistent across databases (Medline and CENTARL) but nevertheless their use was generally appropriate.

An RCT study design filter was used in MEDLINE and EMBASE. The origin of the filter was not stated, so it is not possible to say of it has been tested. However, the ERG conducted some additional searches to check if this might have resulted in relevant studies being missed and, in our opinion it would not have done so.

The manufacturer's submission states that the searching for RCTs was first performed in October 2011, and an update was performed in October 2012. The strategy reported in the appendix of the manufacturer's submission, has a publication limit of 2011 and is presumably the strategy used for the

first set of searching. The ERG assumes that this same strategy was used to do the 2012 update, in which case the ERG considers the search strategy for section 6.1, clinical evidence, fit for purpose.

10.1.2 Search strategy for section 6.8, non-RCT evidence

In addition to the searches for section 6.1, additional searches were carried out to locate non-RCT studies for dimethyl fumarate in RRMS.

The terms used for each search facet were appropriate. The search strategies were structured using a combination of subject indexing and free text search terms and search facets were appropriately combined using Boolean operators. (There was some redundancy in the use of abbreviations for multiple sclerosis in conjunction with the full terms, but this is unlikely to affect the overall performance of the search.) Truncation and wildcards were used appropriately. Appropriately, no study design filter was used in any of the databases.

The search strategy for section 6.8, Non-RCT evidence, was appropriate.

10.1.3 Search strategy for section 6.9, Adverse events

The submission states that the search strategy for section 6.1, clinical evidence, was designed to identify eligible studies for adverse events associated with dimethyl fumarate so a separate search for adverse events evidence was not carried out.

The search strategy for section 6.9, Adverse events, was appropriate.

10.1.4 Search strategy for cost-effectiveness studies (manufacturer's submission, section 7.1 and section 7.2)

The submission gave detailed descriptions of the search strategies and met NICE requirements. All of the databases specified by NICE in the specification for manufacturer/sponsor submission of evidence (MEDLINE, MEDLINE In-Process, EMBASE, EconLIT and NHS EED) were searched.

In addition, reference lists of previous trials and systematic reviews were searched as well as unpublished data from clinical study reports held by the manufacturer.

The following Web sites were searched for conference abstracts that were published from 2009 to 2012: European Committee for the Treatment and Research in Multiple Sclerosis (ECTRIMS), Americas Committee for Treatment and Research in Multiple Sclerosis (ACTRIMS), American Academy of Neurology (AAN), American Neurological Association (ANA), European Federation of Neurological Societies (EFNS), International Society for Pharmacoeconomics and Outcomes Research – US and EU (ISPOR).

The strategies aimed to retrieve cost-effectiveness studies relating to dimethyl fumarate, or its comparators. The terms used for each search facet were generally appropriate (as with the searches for non-RCT evidence, there was some redundancy in the use of abbreviations for multiple sclerosis in conjunction with the full terms, but this is unlikely to affect the overall performance of the search). The search strategies were structured using a combination of subject indexing and free text search terms and search facets were correctly combined using Boolean operators. Truncation and wildcards were used appropriately. The use of an economics filter to search NHS EED is not necessary and may have caused the manufacturers to miss some reports of relevant studies.

10.1.5 Search strategy for measurement and valuation of health effects (manufacturer's submission sections 7.4.5 and 7.4.6)

Searches were conducted for HRQoL data relating to epilepsy or seizure. The databases searched included MEDLINE, MEDLINE In-Process, EMBASE, and NHS EED, as specified by NICE. Despite being a required database, EconLIT was not searched for this section.

Additional searches were carried out on the following resources: European Committee for the Treatment and Research in Multiple Sclerosis (ECTRIMS), Americas Committee for Treatment and Research in Multiple Sclerosis (ACTRIMS), American Academy of Neurology (AAN), International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and European Charcot foundation.

The submission included details of the database service providers used, the date span of the searches, the specific databases searched and the dates when searches were conducted. The search strategy for the bibliographic databases suitably comprised a combination of subject indexing and free text search terms; truncation and wild cards were used appropriately.

While the choice of search terms was appropriate, the combination of search facets differs significantly from the approach used in published research, for example Papaioannou 2013[50] where all HRQoL terms are combined using the Boolean operator "OR". The strategy employed in the manufacturer's report combines some of the HRQoL terms with "AND". This will result in fewer studies being identified. Consequently it remains unclear whether the strategy for the bibliographic databases will have retrieved all relevant studies.

10.1.6 Search strategy for resource identification, measurement and valuation (manufacturer's submission section 7.5.3)

Searches were not conducted for resource use in RRMS.

10.2 Additional ERG analyses details

10.2.1 Drug prices

The manufacturer's proposed PAS price compared to risk sharing scheme prices for Rebif 22µg, Rebif 44µg, Avonex, Betaferon, glatiramer acetate; 35% or 53% discount for fingolimod, and no price reduction for natalizumab. These are presented in Table 61.

Table 61: Discount drug prices

	Prices used
Dimethyl fumarate	[redacted]
Rebif 22µg	7,513
Rebif 44µg	8,942
Avonex	8,502
Glairamer acetate	5,823
Fingolimod (35% reduction)	12,464
Fingolimod (53% reduction)	9,109
Natalizumab	14,690
Betaferon	7,259

Probabilistic sensitivity analyses

Probabilistic sensitivity analyses were run using the prices listed in Table 61. The number of iterations was set to 10,000.

Cell 'PSA!E5'=10,000.

1. Relative importance of different outcomes

1.1 Annualised relapse rate relative to placebo for each comparator set to 1.

'Inputs!G68'=1 and 'Inputs!M68'=1.

1.2 Relative risk of progression relative to placebo for each comparator set to 1.

'Inputs!H68'=1 and 'Inputs!N68'=1.

1.3 No waning effect

'Inputs!J85:Q86'=100%

1.4 Complete waning effect after 2 years

'Inputs!H85:Q86'=0%

1.5 No adverse events

'Inputs!G99:122'=0, and

'Inputs!M99:122'=0

10.2.2 Monitoring costs

The alternative resource assumptions listed in Table 62 along with unit costs stated in Table 63 were used to derive the year 1 and year 2+ monitoring costs for each drug presented in Table 64.

Table 62: Alternative resource assumptions

	Resource use	
Treatment	**Year 1**	**Subsequent years**
Avonex	*2 neurology visits*	1 neurology visits
Rebif 22µg	3 full blood counts	2 full blood counts
Rebif 44µg	3 liver function tests	2 liver function tests
Betaferon	*4 MS Nurse visits (at 1,3,6 and 12 months)*	*2 MS Nurse visits*
Dimethyl Fumarate	3 neurology visits	*2 neurology visits*
	3 full blood counts	2 full blood counts
	3 liver function tests	2 liver function tests
	3 renal function tests	2 renal function tests
	4 MS Nurse visits	*2 MS Nurse visits*
Fingolimod	3 neurology visits	*2 neurology visits*
	3 full blood counts	2 full blood counts
	3 liver function tests	2 liver function tests
	1 basic metabolism test	2 basic metabolism test
	1 ophthalmology visit	*2 MS Nurse visits*
	1 patient observation after first administration by healthcare professional	
	4 MS Nurse visits	
Glatiramer acetate	2 neurology visits	1 neurology visit
	2 full blood counts	*2 MS Nurse visits*
	2 liver function tests	
	4 MS Nurse visits	
Natalizumab	*3 neurology visits*	2 neurology visits
	2 full blood counts	*1 MRI scan*
	2 liver function tests	*2 MS Nurse visits*
	1 MRI scan	

	4MS Nurse visits	

Table 63: resource use unit costs

Resource	Unit cost (£)	Source
Neurology visit	£589.83	NHS Reference costs 2011-12: weighted average cost per day of AA30A and AA30B (day case) (31)
Full blood count	£1.23	NHS Reference costs 2011-12: pathology services (Biochemistry - DAP841) (31)
Liver function	£1.23	NHS Reference costs 2011-12: pathology services (Biochemistry - DAP841) (31)
Basic metabolism	£1.23	NHS Reference costs 2011-12: weighted average cost per day of RA01A, RA01B and RA01C (31)
Renal function test	£1.23	NHS Reference costs 2011-12: pathology services (Biochemistry - DAP841) (31)
MRI scan	£149.49	NHS Reference costs 2011-12: pathology services (Biochemistry - DAP841) (31)
Patient observation after first admin	£538.00	2012-13 tariff - outpatient attendances [WF01B First Attendance - Single Professional] (32)
Ophthalmology visit	£115.00	2012-13 tariff - admitted patient care & outpatient procedures (AA30Z) [Combined day case/ordinary elective tariff] (32)
Nurse visit	£58.00	p.207 Table 69 of submission

Table 64: Monitoring costs estimates

	Total monitoring costs (£)	
	Year 1	Subsequent years
Dimethyl fumarate	2,012.56	1,303.04

Rebif 22µg	1,419.04	710.75
Rebif 44µg	1,419.04	710.75
Avonex	1,419.04	710.75
Glairamer acetate	1,416.58	705.83
Fingolimod	2,661.87	1,303.04
Natalizumab*	2,155.90	1,445.15
Betaferon	1,419.04	710.75

Neurology visits

The resource assumptions in Table 62 were applied as well as the unit costs in Table 63 except for the cost of a neurology visit which was reduced to £205 from £589.83. This resulted in the monitoring costs for each drug listed in Table 65.

Table 65: Monitoring cost estimates with neurology cost visits

	Total monitoring costs (£)	
	Year 1	Subsequent years
Dimethyl fumarate	858.07	533.38
Rebif 22µg	649.38	325.92
Rebif 44µg	649.38	325.92
Avonex	649.38	325.92
Glairamer acetate	646.92	321
Fingolimod	1,507.38	533.38
Natalizumab*	1,001.41	675.49
Betaferon	649.38	325.92

10.2.3 Disease progression

10.2.3.1 EDSS baseline population distribution

Alternative analyses were run with 100% of the population starting in different EDSS states: 0-5.

10.2.3.2 Utility estimates for EDSS states

Utility estimates for the EDSS states for both RRMS and SPMS with or without relapse from Orme et al. were used.[39] These are presented in Table 66.

Table 66: Orme et al utility estimates[39]

Clinical presentation	Disease type	EDSS state									
		0	1	2	3	4	5	6	7	8	9

No relapse	RRMS	0.870	0.799	0.705	0.574	0.610	0.518	0.458	0.297	-0.049	-0.195
	SPMS	0.825	0.754	0.660	0.529	0.565	0.473	0.413	0.252	-0.094	-0.240
Relapse	RRMS	0.799	0.728	0.634	0.503	0.539	0.447	0.387	0.226	-0.120	-0.266
	SPMS	0.754	0.683	0.589	0.458	0.494	0.402	0.342	0.181	-0.165	-0.311

10.2.3.3 Utility estimates from TA127

Utility estimates from the Natalizumab submission for both RRMS and SPMS with or without relapse were used.[6] These are presented in Table 67.

Table 67: Natalizumab utility estimates (TA127)[6]

Clinical presentation	**Disease type**	**EDSS state**									
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
No relapse	RRMS	0.909	0.844	0.745	0.611	0.654	0.558	0.495	0.437	-0.007	-0.151
	SPMS	0.865	0.8	0.701	0.568	0.61	0.514	0.451	0.393	-0.051	-0.195
Relapse	RRMS	0.9	0.835	0.735	0.602	0.645	0.548	0.485	0.427	-0.016	-0.16
	SPMS	0.856	0.791	0.692	0.559	0.601	0.505	0.442	0.384	-0.06	-0.204

10.2.3.4 RRMS and SPMS EDSS state costs from TA127

Alternative costs for the RRMS and SPMS EDSS states were used based on those derived in the Natalizumab.[6] These are presented in Table 68.

Table 68: TA127 EDSS state costs

	Natalizumab (2005- Inflated to 2012 costs)[6]	
EDSS State	RRMS	SPMS
0	785	854
1	1,140	1,209
2	1,086	1,155
3	3,392	3,461
4	2,160	2,229
5	3,128	3,196
6	3,869	3,938
7	9,081	9,150
8	21,363	21,432
9	20,056	20,124

RRMS and SPMS EDSS state costs from TA254

Alternative costs for the RRMS and SPMS EDSS states were used based on those derived from Tyas et al.[46] These are presented in Table 69.

Table 69: TA254 EDSS state costs

	Tyas (2005)- Inflated to 2012 costs[46]	
EDSS State	RRMS	SPMS
0	3,426	3,771
1	4,362	4,707
2	5,691	6,035
3	8,685	9,030
4	5,945	6,290
5	9,534	9,878
6	12,534	16,925
7	21,389	21,733
8	32,100	32,444
9	31,094	31,438

10.2.3.5 Transition rates to SPMS

Sensitivity analysis was conducted on the transition rates to SPMS. The rates were increased by 50% or reduced by 50%. The transition rates to SPMS and one EDSS state higher that were used are presented in Table 70.

Table 70: Transition rate estimates following ERG assumptions

SPMS EDSS	Increased by 50%	Reduced by 50%
1	0.005	0.002
2	0.048	0.021
3	0.175	0.078
4	0.316	0.141
5	0.448	0.200
6	0.356	0.159
7	0.380	0.170
8	0.229	0.102
9	1	1

10.2.3.6 Discontinuation rates

50% after 2 years

Discontinuation rates were assumed to be 50% of the original rate after 2 years.

'Inputs!J81:Q81'=0.5* tys_ae_drop_year2, and

'Inputs!J82:Q82'=0.5*bi_ae_drop_year2

0% after 2 years

Discontinuation rates were assumed to be 0% after 2 years.

'Inputs!J81:Q82'=0%

10.2.3.7 Upper and lower CIs for discontinuation rates

Using the upper and lower confidence intervals of the relative risks for discontinuation due to any cause reported in Figure 30, P151, section 6.7.6, and using the discontinuation risk of 12.55 for dimethyl fumarate quoted in Table 52, P183, section 7.3.1, upper and lower discontinuation risks for each comparator were calculated as presented in Table 71.

Table 71: Discontinuation rates following ERG assumptions

	Discontinuation risk	
	Lower limit	Upper limit
Rebif 22µg	9.77	31.61
Rebif 44µg	11.69	22.10
Avonex	9.08	18.87
Glairamer acetate	8.79	13.75
Fingolimod	7.95	14.18
Natalizumab	7.51	16.73
Betaferon	6.42	12.60

10.2.4 Relapses

10.2.4.1 Relapse rates based on MS survey

Baseline relapse rates were set at the values estimated using the MS survey and reported in Table 47, P178, section 7.3.1, presented here in Table 72.

Table 72: Annual relapse rates for each EDSS states for both RRMS and SPMS from the MS survey

EDSS	Annual relapse rate (RRMS)	Annual relapse rate (SPMS)
0	0.71	0.00
1	0.73	0.00
2	0.68	0.47
3	0.72	0.88
4	0.71	0.55
5	0.59	0.52
6	0.49	0.45
7	0.51	0.34
8	0.51	0.34
9	0.51	0.34

10.2.4.2 Alternative relapse cost assumptions were tested

- £280.41/relapse: £228 from the natalizumab submission (TA127)[6] inflated to 2012 prices
- £1996.09/relapse: £1623 from Tyas et al (2005)[46] inflated to 2012 prices
- £3039/relapse: from the fingolimod submission (TA254)[7] (AA30Z code for 2011/12 NHS Reference Costs)
- £607.8/relapse: 20% of patients assumed admitted to hospital at £3039 per admission and 80% incurring no costs

10.2.5 Adverse events

The incidence of adverse events for the comparators were calculated using the incidence of adverse events for dimethyl fumarate stated in the manufacturer's model, and the relative risks of the adverse events for dimethyl fumarate compared to each comparator derived from the MTC. These were obtained from Table 35, P148-50, section 6.7.6. The relative risks are presented in Table 73.

Table 73 Relative risk of adverse events from manufacturer's submission

	Relative risk of adverse event for dimethyl fumarate compared to the comparator						
	Rebif 22µg	Rebif 44µg	Glatiramer acetate	Avonex	Betaferon	Fingolimod	Natalizumab
Abdominal pain			[redacted]				[redacted]
Abdominal pain upper							
ALT increased	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Arthralgia		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Atrioventricular conduction block			[redacted]	[redacted]		[redacted]	
Back pain		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Bradycardia			[redacted]	[redacted]		[redacted]	
Chest pain			[redacted]			[redacted]	[redacted]

Cough			[redacted]	[redacted]		[redacted]	
Depression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhea			[redacted]	[redacted]		[redacted]	
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Flu-like symptoms	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Flushing			[redacted]				
Gastroenteritis			[redacted]				[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Influenza		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Leucopenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Lower respiratory tract infection			[redacted]			[redacted]	[redacted]
Nausea		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
Pain in extremity			[redacted]	[redacted]	[redacted]	[redacted]	
Pruritus			[redacted]				[redacted]
Rash		[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
Urinary tract infection		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

10.3 Philips checklist

Quality criterion	Question(s)	Response (√, X, or NA)	Comments
S1	Is there a clear statement of the decision problem?	√	
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?	√	
	Is the primary decision-maker specified?	√	NHS and Personal Social Services
S2	Is the perspective of the model stated clearly?	√	
	Are the model inputs consistent with the stated perspective?	√	The model measures the progression of MS through disability progression and regression, relapse rates and adverse events.
	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?	√	
S3	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	√	Appears to be consistent although there is an issue with EDSS fully estimating the disability surrounding MS
	Are the sources of data used to develop the structure of the model specified?	√	Sources of data are specified and appear broadly consistent with previous submissions for DMTs in MS
	Are the causal relationships described by the model structure justified appropriately?	√	
S4	Are the structural assumptions transparent and justified?	√	
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	√	
S5	Is there a clear definition of the options under evaluation?	√	
	Have all feasible and practical options been evaluated?	X	BSC is excluded
	Is there justification for the exclusion of feasible options?	√	Justified as this is not part of the NICE scope
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	√	
S7	Is the time horizon of the model sufficient to reflect all important differences between options?	√	30 years appears sufficient

Quality criterion	Question(s)	Response (√, X, or NA)	Comments
	Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	√	
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?	√	
S9	Is the cycle length defined and justified in terms of the natural history of disease?	√	
D1	Are the data identification methods transparent and appropriate given the objectives of the model?	X	There is some lack of clarity surrounding some of the data sources
	Where choices have been made between data sources, are these justified appropriately?	√	Broadly speaking, yes
	Has particular attention been paid to identifying data for the important parameters in the model?	√	
	Has the quality of the data been assessed appropriately?	√	
	Where expert opinion has been used, are the methods described and justified?	X	Expert opinion was elicited for clinical parameters and model validation but the methods were not described
D2	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	√	
D2a	Is the choice of baseline data described and justified?	√	
	Are transition probabilities calculated appropriately?	√	
	Has a half-cycle correction been applied to both cost and outcome?	√	
	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?	√	
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	√	.
	Have alternative extrapolation assumptions been explored through sensitivity analysis?	√	Weaning effects of treatment were included

Quality criterion	Question(s)	Response (√, X, or NA)	Comments
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	√	Assumptions have been documented but only justified by referring to a previous submission
D2c	Are the costs incorporated into the model justified?	√	
	Has the source for all costs been described?	√	
	Have discount rates been described and justified given the target decision-maker?	√	
D2d	Are the utilities incorporated into the model appropriate?	√	
	Is the source for the utility weights referenced?	√	
	Are the methods of derivation for the utility weights justified?	√	
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	X	UK MS Survey and London Ontario data not sufficiently referenced (but also not in the public domain so they may be non-applicable)
	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	√	
	Is the process of data incorporation transparent?	√	
	If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	X	Probability distributions were not specified for several of the significant parameters, such as treatment waning effect and the annual discontinuation risk
	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	X	Only clear for the parameters where the distributions have been specified
D4	Have the four principal types of uncertainty been addressed?	X	Heterogeneity has not been addressed – no sub-group analysis
	If not, has the omission of particular forms of uncertainty been justified?	X	
D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	X	.
D4b	Is there evidence that structural uncertainties have been	X	.

Quality criterion	**Question(s)**	**Response (√, X, or NA)**	**Comments**
	addressed via sensitivity analysis?		
D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	√	
D4d	Are the methods of assessment of parameter uncertainty appropriate?	X	PSA has been used but some of the distributions used were not specified
	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	X	Stated clearly but the ranges are not justified
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	√	
C2	Are any counterintuitive results from the model explained and justified?	NA	
	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?	NA	
AEs, Adverse Events; HRQoL, Health Related Quality of Life; KOL, Key Opinion Leaders; mCRC, metastatic Colorectal Cancer; MS, Manufacturer's Submission, PD, Progressive Disease: RCT – Randomised Clinical Trial			